JOURNAL OF FUNCTIONAL ANALYSIS 44, 212–257 (1981)

The Malliavin Calculus, a Functional Analytic Approach

DANIEL W. STROOCK*

Department of Mathematics, University of Colorado, Boulder, Colorado 80309

Communicated by the Editors

Received May 1981

INTRODUCTION

In his groundbreaking articles [8, 9], Malliavin introduced a technique for obtaining elliptic regularity results using function space calculus. In his formulation, the function space calculus on which his theory rests is the martingale calculus coming from the Ornstein–Uhlenbeck process on Wiener space (cf. [12] for a detailed discussion of Malliavin's theory from this point of view). Although this formulation has a great deal to recommend it and is particularly pleasing to afficionados of the so-called "Brownian sheet," most analysts (including those who are reasonably facile with the machinery of probability theory) are unlikely to find the technical difficulties inherent in this approach outweighed by the eventual rewards.

Shigekawa [11] discovered that there is an alternative formulation of Malliavin's theory. In Shigekawa's formulation, no mention need be made of the Ornstein–Uhlenbeck process. Instead, what he relies on is a Sobolev type extension of the Frechét dervative with Wiener measure playing the role played by Lebesgue measure in the finite dimensional context. A third formulation of Malliavin's theory was recently provided by Bismut [1]. Bismut's idea is to exploit the quasi-invariance of Wiener measure as it is manifested in a beautiful relation due to Haussmann [4].

The present article has two aims. In the first place, Section 1, 2, and 3 are devoted to yet another formulation of Malliavin's calculus. Although the end results are precisely the same as those obtained in [12], no reference is made to the Ornstein–Uhlenbeck process on which the original theory depended. Instead, the machinery is entirely that of elementary functional analysis and is devoid of any sophisticated martingale calculus. It has been my intent that these three sections should be readily accessible to any analyst who has had some acquintance with Wiener measure.

* Research sponsored in part by NSF Grant MCS 80-07300.

0022-1236/81/140212–46$02.00/0

The second aim of this paper is to whet the reader's appetite by providing him with some applications of Malliavin's calculus. With this goal in mind, I have concentrated on examples for which I know of no other technique that works. All these examples are what might be called "partially elliptic" problems in that they involve operators which are elliptic in some but not all directions. So far as I can tell, such problems defy the more usual P.D.E. techniques because the regularity that one can expect occurs only in the "good" (i.e., the elliptic) directions and the equations are not autonomous unless one deals with all directions simultaneously. Thus these results involve projection onto the "good" directions, and it is for this reason that the Malliavin technique is ideally suited to handle them. Indeed, the basic connection between P.D.E.'s and Wiener space involves identifying the marginal distributions of certain Wiener space functionals as the solutions to certain P.D.E.'s. Malliavin's calculus gives one a way of measuring the regularity of the Wiener space functionals; applications of Malliavin's calculus to P.D.E.'s come from projecting these regularity results in Wiener space down to statements about the marginal distributions. Thus, it is not surprising that Malliavin's technique is well suited to problems in which projection must play a key role.

Further evidence of the power that Malliavin's technique has to handle new problems can be found in the article by Michel [10] and the article [5]. Michel's paper contains an application to a problem coming from filtering theory, whereas [5] treats a situation coming from statistical mechanics. In neither case is it clear that the results obtained could have been derived by more standard P.D.E. methodology.

1. Symmetric Diffusion Semigroups, Generalities

Let (E, F, m) be a separable probability space. We will say that the family of operators $\{T_\tau : \tau \geqslant 0\}$ on $L^2(m)$ is a *symmetric Markov semigroup on* $L^2(m)$ if

$$\begin{aligned}&\text{(i) for each } \tau > 0,\ T_\tau \text{ is a non-negative self-adjoint contraction on } L^2(m) \text{ and } T_\tau 1 = 1;\\ &\text{(ii) } T_{\tau+\sigma} = T_\tau \circ T_\sigma,\quad \tau, \sigma \geqslant 0, \text{ and } T_\tau \to I \text{ strongly as } \tau \searrow 0.\end{aligned} \tag{1.1}$$

It is a simple matter to conclude from (1.1) that for each $1 \leqslant q < \infty$ and $\tau > 0$, T_τ determines a unique non-negative contraction $T_\tau^{(q)}$ on $L^q(m)$. Moreover, if $1 \leqslant q < \infty$, then it is easy to check that $\{T_\tau^{(q)} : \tau > 0\}$ is a strongly continuous semigroup on $L^q(m)$. For $1 \leqslant q < \infty$, let $A^{(q)}$ denote the generator of $\{T_\tau^{(q)} : \tau > 0\}$.

(1.2) LEMMA. *For all* $1 < q < \infty$, $A^{(1)}$ *is an extention of* $A^{(q)}$. *Moreover,* $\Phi \in \mathrm{Dom}(A^{(q)})$ *if and only if* $\Phi \in L^q(m) \cap \mathrm{Dom}(A^{(1)})$ and $A^{(1)}\Phi \in L^q(m)$.

Proof. The first part is an immediate consequence of the fact that $T_\tau^{(1)}$ is an extention of $T_\tau^{(q)}$ for all $\tau > 0$ and $1 < q < \infty$. Moreover the "only if" half of the second part is obvious from $A^{(q)} \subseteq A^{(1)}$. Finally, to prove the "if" half, suppose that $\Phi \in L^q(m) \cap \mathrm{Dom}(A^{(1)})$ and that $A^{(1)}\Phi \in L^q(m)$. Then

$$\frac{1}{\tau}(T_\tau^{(q)}\Phi - \Phi) = \frac{1}{\tau}(T_\tau^{(1)}\Phi - \Phi) = \frac{1}{\tau}\int_0^\tau T_\sigma^{(1)}A^{(1)}\Phi\, d\sigma$$
$$= \frac{1}{\tau}\int_0^\tau T_\sigma^{(q)}A^{(1)}\Phi\, d\sigma \xrightarrow{L^q(m)} A^{(1)}\Phi$$

as $\tau \searrow 0$. Q.E.D.

We will say that the quadruple $(\mathscr{L}, T., \mathscr{D}, m)$ is a *symetric diffusion semigroup on* $L^2(m)$ *with generator* $\mathscr{L}$ if $\{T_\tau : \tau > 0\}$ is a symmetric Markov semigroup on $L^2(m)$, $\mathscr{L} = A^{(2)}$, and $\mathscr{D} \subseteq \bigcap_{1 \leqslant q < \infty} A^{(q)}$ is an algebra with the properties that

(i) graph $(\mathscr{L}|_{\mathscr{D}})$ is dense in graph $(\mathscr{L})$,

(ii) for all $\Phi \in \mathscr{D}$ and $F \in \mathscr{S}(R^1)$, $F \circ \Phi \in \mathrm{Dom}(\mathscr{L})$ and $\mathscr{L}\Phi = \frac{1}{2}\langle \Phi, \Phi\rangle_{\mathscr{L}} F'' \circ \Phi + \mathscr{L}\Phi F' \circ \Phi$, (1.3)

where

$$\langle \Phi, \Psi\rangle_{\mathscr{L}} \equiv \mathscr{L}(\Phi \cdot \Psi) - \Phi\mathscr{L}\Psi - \Psi\mathscr{L}\Phi, \qquad \Phi, \Psi \in \mathscr{D}. \tag{1.4}$$

(1.5) LEMMA. *Let* $(\mathscr{L}, T., \mathscr{D}, m)$ *be a symmetric diffusion semigroup. Then* $\langle \Phi, \Phi\rangle_{\mathscr{L}} \geqslant 0$ *(a.s., m) for all* $\Phi \in \mathscr{D}$. *Moreover, if* $\Phi, \Psi \in \mathscr{D}$, *then* $|\langle \Phi, \Psi\rangle|_{\mathscr{L}} \leqslant \langle \Phi\rangle_{\mathscr{L}}\langle \Psi\rangle_{\mathscr{L}}$ *(a.s., m) and* $\langle \Phi + \Psi\rangle_{\mathscr{L}} \leqslant \langle \Phi\rangle_{\mathscr{L}} + \langle \Psi\rangle_{\mathscr{L}}$ *(a.s., m), where*

$$\langle \Phi\rangle_{\mathscr{L}} \equiv (\langle \Phi, \Phi\rangle_{\mathscr{L}})^{1/2}, \qquad \Phi \in \mathscr{D}. \tag{1.6}$$

Finally, if $\Phi, \Psi \in \mathscr{D}$, *then*

$$E^m[\langle \Phi, \Psi\rangle_{\mathscr{L}}] = -2E^m[\Phi\mathscr{L}\Psi]. \tag{1.7}$$

Proof. First note that if $\Phi \in \mathscr{D}$, then $T_\tau(\Phi^2) \geqslant (T_\tau\Phi)^2$ (a.s., m) for all $\tau \geqslant 0$. To see this, let $\Psi \in L^2(m)$ be a non-negative function satisfying $E^m[\Psi] = 1$. Then for each $\tau > 0$, $T_\tau\Psi$ has the same properties. Hence by Jensen's inequality:

$$(T_\tau\Phi^2, \Psi)_{L^2(m)} = (\Phi^2, T_\tau\Psi)_{L^2(m)} \geqslant (\Phi, T_\tau\Psi)^2_{L^2(m)} = (T_\tau\Phi, \Psi)^2_{L^2(m)}.$$

Since this is true for all nonnegative $\Psi \in L^2(m)$ with $E^m[\Psi]=1$, it follows that $T_\tau \Phi^2 \geqslant (T_\tau \Phi)^2$ (a.s., m).

We now can prove that $\langle \Phi, \Phi\rangle_{\mathscr{L}} \geqslant 0$ (a.s., m). Indeed, given a nonnegative $\Psi \in L^2(m)$, note that

$$\begin{aligned}(\langle \Phi, \Phi\rangle_{\mathscr{L}}, \Psi)_{L^2(m)} &= \lim_{\tau \searrow 0} 1/\tau((T_\tau \Phi^2 - \Phi^2) - 2\Phi(T_\tau \Phi - \Phi), \Psi)_{L^2(m)} \\ &\geqslant \overline{\lim_{\tau\searrow 0}}\ 1/\tau((T_\tau \Phi)^2 - 2\Phi T_\tau \Phi + \Phi^2, \Psi)_{L^2(m)} \\ &= \overline{\lim_{\tau\searrow 0}}\ 1/\tau((T_\tau \Phi - \Phi)^2, \Psi)_{L^2(m)} \geqslant 0.\end{aligned}$$

Thus $\langle \Phi, \Phi\rangle_{\mathscr{L}} \geqslant 0$ (a.s., m).

The proof that $|\langle \Phi, \Psi\rangle_{\mathscr{L}}| \leqslant \langle \Phi\rangle_{\mathscr{L}} \langle \Psi\rangle_{\mathscr{L}}$ (a.s., m) is now just like the derivation of the usual Schwartz inequality. That is, for $\lambda > 0$:

$$\begin{aligned}&\lambda^2 \langle \Phi\rangle^2_{\mathscr{L}} \pm 2\langle \Phi, \Psi\rangle_{\mathscr{L}} + 1/\lambda^2 \langle \Psi\rangle^2_{\mathscr{L}} \\ &\quad = \langle \lambda\Phi \pm 1/\lambda \Psi, \lambda\Phi \pm 1/\lambda\Psi\rangle_{\mathscr{L}} \geqslant 0 \qquad \text{(a.s., } m).\end{aligned}$$

Thus there is one m-null set N such that:

$$|\langle \Phi, \Psi\rangle_{\mathscr{L}}| \leqslant 1/2(\lambda^2 \langle \Phi\rangle^2_{\mathscr{L}} + 1/\lambda^2 \langle \Psi\rangle^2_{\mathscr{L}}), \qquad \lambda > 0,$$

off of N. Given a point not in N at which $\langle \Phi\rangle_{\mathscr{L}} \cdot \langle \Psi\rangle_{\mathscr{L}} = 0$, it is clear that $\langle \Phi, \Psi\rangle_{\mathscr{L}} = 0$. On the other hand, if at a point not in N one has $\langle \Phi\rangle_{\mathscr{L}} \cdot \langle \Psi\rangle_{\mathscr{L}} > 0$, simply take $\lambda^2 = \langle \Psi\rangle_{\mathscr{L}} / \langle \Phi\rangle_{\mathscr{L}}$.

The inequality $\langle \Phi + \Psi\rangle_{\mathscr{L}} \leqslant \langle \Phi\rangle_{\mathscr{L}} + \langle \Psi\rangle_{\mathscr{L}}$ (a.s., m) is an immediate consequence of the preceding.

Finally, to prove (1.7) simply observe that

$$\begin{aligned}&E^m[\langle \Phi, \Psi\rangle_{\mathscr{L}}] = E^m[\mathscr{L}(\Phi \cdot \Psi)] - E^m[\Phi \mathscr{L} \Psi] - E^m[\Psi \mathscr{L} \Phi], \\ &E^m[\mathscr{L}(\Phi \cdot \Psi)] = E^m[(\Phi \cdot \Psi)\, \mathscr{L} 1] = 0,\end{aligned}$$

and

$$E^m[\Phi \mathscr{L} \Psi] = E^m[\Psi \mathscr{L} \Phi]. \qquad \text{Q.E.D.}$$

(1.8) LEMMA. *The bilinear map* $\langle \cdot\,.\ \cdot \rangle_{\mathscr{L}}$ *on* $\mathscr{D} \times \mathscr{D}$ *into* $L^2(m)$ *admits a unique extention as a graph*($\mathscr{L}$)*-continuous bilinear map of* $\mathrm{Dom}(\mathscr{L}) \times \mathrm{Dom}(\mathscr{L})$ *into* $L^1(m)$. *Moreover, if* $n \geqslant 1$, $\Phi = (\Phi_1, \ldots, \Phi_n) \in (\mathrm{Dom}(\mathscr{L}))^n$, *and* $F \in C_b^2(R^n)$, *then* $F \circ \Phi \in \mathrm{Dom}(A^{(1)})$ *and*

$$A^{(1)}(F \circ \Phi) = 1/2 \sum_{k,l=1}^{n} \langle \Phi_k, \Phi_l\rangle_{\mathscr{L}} \frac{\partial^2 F}{\partial x_k \partial x_l} \circ \Phi + \sum_{k=1}^{n} \mathscr{L}\Phi_k \frac{\partial F}{\partial x_k} \circ \Phi. \tag{1.9}$$

Proof. We will have shown that the desired unique extention exists once we show that for $\{\Phi_n\}_1^\infty \subseteq \mathscr{D}$ satisfying $\|\Phi_n - \Phi_m\|_{L^2(m)} + \|\mathscr{L}\Phi_n - \mathscr{L}\Phi_m\|_{L^2(m)} \to 0$: $\|\langle \Phi_n, \Phi_n\rangle_{\mathscr{L}} - \langle \Phi_m, \Phi_m\rangle_{\mathscr{L}}\|_{L^1(m)} \to 0$. To this end, note that by (1.7)

$$\sup_n \|\langle \Phi_n, \Phi_n\rangle_{\mathscr{L}}\|_{L^1(m)} \leqslant 2 \sup_n \|\Phi_n\|_{L^2(m)} \|\mathscr{L}\Phi_n\|_{L^2(m)} < \infty$$

and that

$$\|\langle \Phi_n - \Phi_m, \Phi_n - \Phi_n\rangle_{\mathscr{L}}\|_{L^1(m)} \leqslant 2\|\Phi_n - \Phi_m\|_{L^2(m)} \|\mathscr{L}\Phi_n - \mathscr{L}\Phi_m\|_{L^2(m)} \to 0.$$

Thus

$$\begin{aligned}\|\langle \Phi_n, \Phi_n\rangle_{\mathscr{L}} - \langle \Phi_m, \Phi_m\rangle_{\mathscr{L}}\|_{L^1(m)} &= \|(\langle \Phi_n\rangle_{\mathscr{L}} - \langle \Phi_m\rangle_{\mathscr{L}})(\langle \Phi_n\rangle_{\mathscr{L}} + \langle \Phi_m\rangle_{\mathscr{L}})\|_{L^1(m)} \\ &\leqslant (\|\langle \Phi_n\rangle_{\mathscr{L}}\|_{L^2(m)} + \|\langle \Phi_m\rangle_{\mathscr{L}}\|_{L^2(m)}) \|\langle \Phi_n - \Phi_m\rangle_{\mathscr{L}}\|_{L^2(m)} \to 0.\end{aligned}$$

To prove (1.9), first observe that the class of F's for which (1.9) holds is closed under $C_b^2(R^n)$-convergence. Thus it suffices for us to prove (1.9) for F's of the form

$$F(x) = \sum_0^N a_m \cos(\theta_m \cdot x),$$

where $N \geqslant 1$, $\{a_m\}_0^N \subseteq R^1$, and $\{\theta_m\}_0^N \subseteq R^n$. Since (1.9) is clearly linear in F, we now see that it is enough to prove (1.9) where $F(x) = f(\theta \cdot x)$ for some $f \in C_b^2(R^n)$ and $\theta \in R^n$. But assuming that $f \in \mathscr{S}(R^1)$ and that $\Phi \in \mathscr{D}^n$, one sees that (1.9) is immediate from (1.3); and therefore, since $A^{(1)}$ is closed, (1.9) continues to hold for all $f \in C_b^2(R^n)$ and $\Phi \in (\mathrm{Dom}(\mathscr{L}))^n$. Q.E.D.

Given a symmetric diffusion semigroup $(\mathscr{L}, T., \mathscr{D}, m)$ and $2 \leqslant q < \infty$, set $\mathscr{K}_{(q)}(\mathscr{L}) = \{\Phi \in \mathrm{Dom}(A^{(q)}): \langle \Phi\rangle_{\mathscr{L}} \in L^q(m)\}$ and define

$$\|\Phi\|_{\mathscr{K}_{(q)}(\mathscr{L})} = E^m[(\Phi^2 + (\mathscr{L}\Phi)^2 + \langle \Phi\rangle_{\mathscr{L}}^2)^{q/2})^{1/q}. \tag{1.10}$$

Because $A^{(q)}$ is closed, it is clear that $(\mathscr{K}_{(q)}(\mathscr{L}), \|\cdot\|_{\mathscr{K}_{(q)}(\mathscr{L})})$ is a Banach space for each $q \in [2, \infty)$. (Note that $\mathscr{K}_{(2)}(\mathscr{L}) = \mathrm{Dom}(\mathscr{L})$ and that $\|\cdot\|_{\mathscr{K}_{(2)}(\mathscr{L})}$ is equivalent to the graph $(\mathscr{L})$-norm. Thus $(\mathscr{K}_{(2)}(\mathscr{L}), \|\cdot\|_{\mathscr{K}_{(2)}(\mathscr{L})})$ is equivalent to a Hilbert space.) Finally, set $\mathscr{K}(\mathscr{L}) = \bigcap_{1 \leqslant q < \infty} \mathscr{K}_{(q)}(\mathscr{L})$. Clearly, $\mathscr{K}(\mathscr{L})$ can be turned into a countable normed Frechét space. As the next result shows, $\mathscr{K}(\mathscr{L})$ is also an algebra.

(1.11) LEMMA. *Let $\Phi = (\Phi_1, \ldots, \Phi_n) \in (\mathscr{K}(\mathscr{L}))^n$ and $F \in C_{\nearrow}^2(R^n)$ (the space of $F \in C^2(R^n)$ such that $D^\alpha F$ is slowly increasing for all $|\alpha| \leqslant 2$). Then*

$F \circ \Phi \in \mathscr{K}(\mathscr{L})$ and $\mathscr{L}(F \circ \Phi)$ is given by the right hand side of (1.9). *Moreover, for any $\Psi \in \mathscr{K}(\mathscr{L})$,*

$$\langle F \circ \Phi, \Psi \rangle_{\mathscr{L}} = \sum_1^N \langle \Phi_k, \Psi \rangle_{\mathscr{L}} \frac{\partial F}{\partial x_k} \circ \Phi. \tag{1.12}$$

Proof. In view of Lemma (1.2) and the fact that $C^2_{\nearrow}(R^n)$ is an algebra, we will know that $F \circ \Phi \in \mathscr{K}(\mathscr{L})$ once we show that $F \circ \Phi \in \mathrm{Dom}(A^{(1)})$ and that $A^{(1)}(F \circ \Phi)$ is given by the right hand side of (1.9). However, we already know $F \circ \Phi \in \mathrm{Dom}(A^{(1)})$ and (1.9) for $F \in C_b^2(R^n)$, and so the desired conclusion follows for all $F \in C^2_{\nearrow}(R^n)$ after an easy limit argument.

The proof of (1.12) follows immediately from the preceding by considering $\tilde{\Phi} = (\Phi_1, ..., \Phi_n, \Psi)$, $\tilde{F}(x, y) = F(x)\, y$, and using the right hand side of (1.9) to compute $\mathscr{L}(\tilde{F} \circ \tilde{\Phi})$. Q.E.D.

(1.13) Lemma. *Let $\Phi \in \mathscr{K}(\mathscr{L})$ be positive and suppose that $1/\Phi \in \bigcap_{1 \leqslant q < \infty} L^q(m)$. Then $1/\Phi \in \mathscr{K}(\mathscr{L})$, $\mathscr{L}(1/\Phi) = 2\langle \Phi, \Phi \rangle_{\mathscr{L}}/\Phi^3 - \mathscr{L}\Phi/\Phi^2$, and $\langle 1/\Phi, \Psi \rangle_{\mathscr{L}} = -1/\Phi^2 \langle \Phi, \Psi \rangle_{\mathscr{L}}$ for $\Psi \in \mathscr{K}(\mathscr{L})$.*

Proof. Since Φ^2 satisfies the hypotheses if Φ does, we will know that $1/\Phi \in \mathscr{K}(\mathscr{L})$ once we show that $1/\Phi \in \mathrm{Dom}(A^{(1)})$ and that $A^{(1)}(1/\Phi) = 2\langle \Phi, \Phi \rangle_{\mathscr{L}}/\Phi^3 - \mathscr{L}\Phi/\Phi^2$. But this is easily shown by considering $F_\varepsilon \circ \Phi$, where $F_\varepsilon(x) = (\varepsilon^2 + x^2)^{-1/2}$ and letting $\varepsilon \searrow 0$. At the same time one can see that $\langle F_\varepsilon \circ \Phi, \Psi \rangle_{\mathscr{L}} \to -1/\Phi^2 \langle \Phi, \Psi \rangle_{\mathscr{L}}$ in $L^1(m)$ and therefore that $\langle -1/\Phi, \Psi \rangle_{\mathscr{L}} = -1/\Phi^2 \langle \Phi, \Phi \rangle_{\mathscr{L}}$. Q.E.D.

(1.14) Theorem. *Let $(\mathscr{L}, T., \mathscr{D}, m)$ b a symmetric diffusion semigroup. Suppose that $\Phi = (\Phi_1, ..., \Phi_n) \in (\mathscr{K}(\mathscr{L}))^n$ and set $A = ((\langle \Phi_k, \Phi_l \rangle_{\mathscr{L}}))_{1 \leqslant k,l \leqslant n}$. Then A is (a.s., m) a non-negative definite symmetric matrix. Moreover, if $A \in (\mathscr{K}(\mathscr{L}))^{n^2}$, then for all $F \in C^1(R^n)$, $\Psi \in \mathscr{K}(\mathscr{L})$, and $1 \leqslant k \leqslant n$:*

$$E^m\left[\left(\frac{\partial F}{\partial x_k} \circ \Phi\right)(\Delta \cdot \Psi)\right] = -[(F \circ \Phi)(\mathscr{H}_k \Psi)\Big], \tag{1.15}$$

where $\Delta = \det A$ and

$$\mathscr{H}_k \Psi = \sum_l (\langle \Phi_l, A^{(k,l)} \Psi \rangle_{\mathscr{L}} + 2(\mathscr{L}\Phi_l)(A^{k,l)}\Psi)) \tag{1.16}$$

with $A^{(k,l)}$ denoting the (k, l)th cofactor of the matrix A. In particular, if $1/\Delta \in \bigcap_{1 \leqslant q < \infty} L^q(m)$, then:

$$E^m\left[\left(\frac{\partial F}{\partial x_k} \circ \Phi\right)\Psi\right] = -.E^m[(F \circ \Psi)(\mathscr{H}_k(\Psi/\Delta))] \tag{1.17}$$

for all $\Psi \in \mathscr{K}(\mathscr{L})$.

Proof. It is clear that A is symmetric. Moreover, for any $\theta \in R^n$, $(\theta, A\theta) = \langle \sum_1^n \theta_k \Phi_k, \sum_1^n \theta_l \Phi_l \rangle_{\mathscr{L}} \geqslant 0$ (a.s., m). Thus there is an m-null set N such that for all $\theta \in R^n$, $(\theta, A\theta) \geqslant 0$ off of N. That is, A is m-a.s. non-negative definite.

To prove (1.15), note that

$$\langle F \circ \Phi, \Phi_l \rangle_{\mathscr{L}} = \sum_{k'} \left(\frac{\partial F}{\partial x_{k'}} \circ \Phi \right) A_{k'l}, \qquad 1 \leqslant l \leqslant n.$$

Thus, by Cramer's rule,

$$\left(\frac{\partial F}{\partial x_k} \circ \Phi \right) \Delta = \sum_l \sum_{k'} \left(\frac{\partial F}{\partial x_{k'}} \circ \Phi \right) A_{k'l} A^{(k,l)} = \sum_l \langle F \circ \Phi, \Phi_l \rangle_{\mathscr{L}} A^{(lk)}.$$

Therefore

$$E^m \left[\left(\frac{\partial F}{\partial x_k} \circ \Phi \right) (\Delta \Psi) \right] = \sum_l E^m [\langle F \circ \Phi, \Phi_l \rangle_{\mathscr{L}} (A^{(lk)} \Psi)].$$

Observing that

$$\begin{aligned} &E^m[\langle F \circ \Phi, \Phi_l \rangle_{\mathscr{L}} (A^{(k,l)} \Psi)] \\ &\quad = E^m[(F \circ \Phi)(\Phi_l \mathscr{L}(A^{(k,l)} \Psi) - A^{(k,l)} \Psi \mathscr{L} \Phi_l - \mathscr{L}(\Phi_l A^{(k,l)} \Psi))] \\ &\quad = -E^m[(F \circ \Phi)(\langle \Phi_l, A^{k,l)} \Psi \rangle_{\mathscr{L}} + 2(\mathscr{L} \Phi_l)(A^{k,l)} \Psi))], \end{aligned}$$

we arrive at (1.15). Q.E.D.

(1.18) LEMMA. *Let μ be a probability measure on R^n and let $1 \leqslant q < \infty$. Suppose that for each $1 \leqslant k \leqslant n$ there is a $\psi_k \in L^q(\mu)$ such that $\int_{R^n} (\partial F / \partial x_k)(x) \, \mu(dx) = -\int_{R^n} F(x) \, \psi_k(x) \, \mu(dx)$ for all $F \in C_0^\infty(R^n)$. Then μ is absolutely continuous with respect to Lebesgue measure on R^n. Moreover, if $q > n$ and $f = d\mu/dx$, then $f \in \hat{C}(R^n)$ (the space of continuous functions which tend to 0 at ∞), there is a $C_q < \infty$ such that $\|f\|_u \leqslant C_q (\sum_1^n \|\psi_k\|_{L^q(\mu)})^n$, and $\|D_k f\|_{L^q(R^n)} \leqslant \|f\|_u^{1 - 1/q} \|\psi_k\|_{L^q(\mu)}$. In particular, if $q > n$, then there is a $C'_q < \infty$ such that $|f(x+k) - f((x)| \leqslant C'_q (\sum_1^n \|\psi_k\|_{L^q(\mu)})^{n(1-1/q)} (1 + \sum_1^n \|\psi_k\|_{L^q(\mu)}) \, |h|^{1-n/q}$, $|h| \leqslant 1$.*

Proof. For $\lambda > 0$ define

$$G_\lambda(x) = \int_0^\infty \frac{e^{-\lambda t}}{(4\pi t)^{n/2}} e^{-|x|^2/4t}, \qquad x \in R^n \backslash \{0\}.$$

Define the measures ν_k, $1 \leqslant k \leqslant n$, by $d\nu_k = \psi_k d\mu$. As tempered distributions, $D_k \mu = \nu_k$. Thus

$$\mu = G_1 * \mu + \sum_{k=1}^{n} (D_k G_1) * \nu_k.$$

Since $G_1 \in L^1(R^n)$ and $D_k G_1 \in L^1(R^n)$, $1 \leqslant k \leqslant n$, it follows that μ is absolutely continuous with respect to Lebesgue measure. Moreover, if $f = d\mu/dx$, then as a tempered distribution: $D_k f = \psi_k f$.

Next suppose that $q > n$ and set $v = f^{1/q}$. Then as a tempered distribution:

$$D_k v = 1/q \psi_k f^{1/q}. \tag{1.19}$$

To see this, set $v_\epsilon = (f + \varepsilon)^{1/q}$. Then it is easy to check that for $\varepsilon > 0$ $D_k v_\epsilon = 1/q(f+\varepsilon)^{1/q-1} D_k f = 1/q(f+\varepsilon)^{1/q-1} \psi_k f$. Clearly $D_k v_\epsilon \to 1/q\psi_k f^{1/q}$ a.e. and $|D_k v_\epsilon| \leqslant 1/q\,|\psi_k| f^{1/q}$. Thus, by Lebesgue's dominated convergence theorem, $D_k v_\epsilon \to 1/q\psi_k f^{1/q}$ in $L^q(R^n)$. Since $v_\epsilon \to v$ in $\mathscr{S}'(R^n)$, we have now proved (1.19). From (1.19), we now see that

$$f^{1/q} = \lambda G_\lambda * (f^{1/q}) + 1/q \sum_{1}^{n} (D_k G_\lambda) * (\psi_k f^{1/q}) \tag{1.20}$$

for every $\lambda > 0$. But if $1/q' = 1 - 1/q$, then, since $q > n$, an easy computation proves that there are $A_q < \infty$ and $B_q < \infty$ such that $\|\lambda G_\lambda\|_{L^{q'}(R^n)} = A_q \lambda^{n/2q}$ and $\|D_k G_\lambda\|_{L^{q'}(R^n)} = B_q \lambda^{n/2q - 1/2}$. In particular, it follows from (1.20) that $f \in \hat{C}(R^n)$ and that

$$\|f\|_u^{1/q} \leqslant \lambda^{n/2q}\left(A_q + B_q \lambda^{-1/2} \sum_{1}^{n} \|\psi_k\|_{L^q(\mu)}\right).$$

for all $\lambda > 0$. Taking $\lambda^{1/2} = B_p \sum_1^n \|\psi_k\|_{L^q(\mu)}/A_q$, we obtain

$$\|f\|_u^{1/q} \leqslant 2A_q(B_q/A_q)^{n/q} \left(\sum_{1}^{n} \|\psi_k\|_{L^q(\mu)}\right)^{n/q};$$

and so $\|f\|_u \leqslant C_q \left(\sum_1^n \|\psi_k\|_{L^q(\mu)}\right)^n$ with $C_q = (2A_q)^q (B_q/A_q)^n$.

Finally (still assuming that $q > n$), we have $D_k f = \psi_k f$ and therefore that

$$\|D_k f\|_{L^q(R^n)} = \left(\int_{R^n} |\psi_k(x)|^q f^q(x)\, dx\right)^{1/q} \leqslant \|f\|_u^{1-1/q} \|\psi_k\|_{L^q(\mu)}.$$

Also,

$$\|f\|_{L^q(R^n)} \leqslant \|f\|_u^{1-1/q}.$$

Thus f is an element of the Sobolev space $\mathscr{W}_q^{(1)}(R^n)$ and $\|f\|_{\mathscr{W}_q^{(1)}(R^n)} \leqslant \|f\|_u^{1-1/q}(1 + \sum_1^n \|\psi_k\|_{L^q(\mu)})$. The desired Hölder estimate now follows from the standard Sobolev embedding theory. Q.E.D.

(1.21) *Remark.* Let everything be as in Theorem (1.14) and set $\mu = m \circ \Phi^{-1}$. Assuming that $1/\Delta \in \bigcap_{1 \leqslant q < \infty} L^q(m)$ and setting $\Psi_k = \mathscr{H}_k(1/\Delta)$, we have

$$E^m \left[\frac{\partial F}{\partial x_k} \circ \Phi \right] = -E^m[(F \circ \Phi)\, \Psi_k], \qquad 1 \leqslant k \leqslant n, \tag{1.22}$$

for all $F \in C_0^\infty(R^n)$. Now define ν_k on R^n by $\nu_k = (\Psi_k m) \circ \Phi^{-1}$, (where $\Psi_k m$ is the measure on (E, F) which is absolutely continuous with respect to m and has Radon–Nikodym derivative Ψ_k). Clearly $\nu_k \ll \mu$ and if $\psi_k = d\nu_k/d\mu$ then $\psi_k \circ \Phi = E^m[\Psi_k \mid \Phi^{-1}(\mathscr{B}_{R^n})]$. Hence (1.22) becomes

$$\int \frac{\partial F}{\partial x_k}(x)\, \mu(dx) = -\int F(x)\, \psi_k(x)\, \mu(dx) \tag{1.23}$$

and

$$\| \psi_k \|_{L^q(\mu)} \leqslant \| \Psi_k \|_{L^q(\mu)}. \tag{1.24}$$

We can therefore apply Lemma (1.18) to conclude that μ admits a density $f \in \hat{C}(R^n)$ with respect to Lebesgue measure and that f is Hölder continuous of every order strictly less than 1.

2. Ornstein–Uhlenbeck Semigroups

The purpose of this section is to develop the machinery with which we will construct in the next section a symmetric diffusion semigroup on Wiener space.

In the next theorem we will be dealing with the following situation. $(E^{(1)}, \mathscr{F}^{(1)}, m^{(1)})$ and $(E^{(2)}, \mathscr{F}^{(2)}, m^{(2)})$ are two separable probability spaces, and $(\mathscr{L}^{(1)}, T_\cdot^{(1)}, \mathscr{D}^{(1)}, m^{(1)})$ and $(\mathscr{L}^{(2)}, T_\cdot^{(2)}, \mathscr{D}^{(2)}, m^{(2)})$ are symmetric diffusion semigroups on $L^2(m^{(1)})$ and $L^2(m^{(2)})$, respectively. $E = E^{(1)} \times E^{(2)}$, $\mathscr{F} = \mathscr{F}^{(1)} \times \mathscr{F}^{(2)}$, $m = m^{(1)} \times m^{(2)}$, $T_\tau = T_\tau^{(1)} \otimes T_\tau^{(2)}$, $\tau > 0$, on $L^2(m)$, and $\mathscr{D} = \operatorname{span}\{\Phi^{(1)} \otimes \Phi^{(2)} \colon \Phi^{(1)} \in \mathscr{D}^{(1)} \text{ and } \Phi^{(2)} \in \mathscr{D}^{(2)}\}$.

(2.1) Theorem. *$\{T_\tau : \tau > 0\}$ is a symmetric Markov semigroup on $L^2(m)$. In fact, if $\mathscr{L}$ is the generator of $\{T_\tau : \tau > 0\}$, then $(\mathscr{L}, T_\cdot, \mathscr{D}, m)$ is a symmetric diffusion semigroup on $L^2(m)$. Finally, if $\Phi^{(1)} \in \operatorname{Dom}(\mathscr{L}^{(1)})$ and $\Phi^{(2)} \in \operatorname{Dom}(\mathscr{L}^{(2)})$, then $\Phi^{(1)} \otimes \Phi^{(2)} \in \operatorname{Dom}(\mathscr{L})$, $\mathscr{L}(\Phi^{(1)} \otimes \Phi^{(2)}) = \Phi^{(1)} \otimes (\mathscr{L}^{(2)}\Phi^{(2)}) + (\mathscr{L}^{(1)}\Phi^{(1)}) \otimes \Phi^{(2)}$, and $\langle \Phi^{(1)} \otimes \Phi^{(2)}, \Phi^{(1)} \otimes \Phi^{(2)} \rangle_{\mathscr{L}} = (\Phi^{(1)})^2 \otimes \langle \Phi^{(2)}, \Phi^{(2)} \rangle_{\mathscr{L}^{(2)}} + \langle \Phi^{(1)}, \Phi^{(1)} \rangle_{\mathscr{L}^{(1)}} \otimes (\Phi^{(2)})^2$. In particular, $\Phi^{(1)} \in \mathscr{K}_{(q)}(\mathscr{L}^{(1)})(\mathscr{K}(\mathscr{L}^{(1)}))$ and $\Phi^{(2)} \in \mathscr{K}_{(q)}(\mathscr{L}^{(2)})(\mathscr{K}(\mathscr{L}^{(2)}))$ implies that $\Phi^{(1)} \otimes \Phi^{(2)} \in \mathscr{K}_{(q)}(\mathscr{L})(\mathscr{K}(\mathscr{L}))$.*

Proof. The following facts are immediate consequences of the theory of semigroups plus the general properties of tensor products: $\{T_\tau : \tau > 0\}$ is a symmetric Markov semigroup on $L^2(m)$ and $\mathrm{span}\{(\Phi^{(1)} \otimes \Phi^{(2)}, \Phi^{(1)} \otimes (\mathscr{L}^{(2)}\Phi^{(2)}) + (\mathscr{L}^{(1)}\Phi^{(1)}) \otimes \Phi^{(2)} : \Phi^{(1)} \in \mathrm{Dom}(\mathscr{L}^{(1)})$ and $\Phi^{(2)} \in \mathrm{Dom}(\mathscr{L}^{(2)})\}$ is a dense subset of graph $(\mathscr{L})$. In particular, graph $(\mathscr{L}|_{\mathscr{D}})$ is dense in graph $(\mathscr{L})$ and $\langle \Phi^{(1)} \otimes \Phi^{(2)}, \Phi^{(1)} \otimes \Phi^{(2)} \rangle_{\mathscr{L}} = (\Phi^{(1)})^2 \otimes \langle \Phi^{(2)}, \Phi^{(2)} \rangle_{\mathscr{L}^{(2)}} + \langle \Phi^{(1)}, \Phi^{(1)} \rangle_{\mathscr{L}^{(1)}} \otimes \Phi^{(2)}$ for $\Phi^{(1)} \in \mathscr{D}^{(1)}$ and $\Phi^{(2)} \in \mathscr{D}^{(2)}$. Thus we will be done once we show that for $F \in \mathscr{S}(R^1)$, $\Phi^{(1)} \in \mathscr{D}^{(1)}$, and $\Phi^{(2)} \in \mathscr{D}^{(2)}$: $F \circ (\Phi^{(1)} \otimes \Phi^{(2)}) \in \mathrm{Dom}(\mathscr{L})$ and $\mathscr{L}(F \circ \Phi)$ is given by the expression in (1.3)(ii).

Let $\eta \in C_0^\infty(R^1)$ be a function satisfying $\eta \equiv 1$ on $[-1, 1]$ and $\eta \equiv 0$ off of $(-2, 2)$. For $M > 0$, define $\bar{\eta}_M(x) = \int_0^x \eta(\xi/M)\, d\xi$. Given $\Phi^{(1)} \in \mathscr{D}^{(1)}$ and $\Phi^{(2)} \in \mathscr{D}^{(2)}$, set $\Phi_M^{(1)} = \eta_M \circ \Phi^{(1)}$, $\Phi_M^{(2)} = \eta_M \circ \Phi^{(2)}$, and $\Phi_M = \Phi_M^{(1)} \otimes \Phi_M^{(2)}$. Clearly $\Phi_M^{(1)} \in \mathscr{K}(\mathscr{L}^{(1)})$, $\Phi_M^{(2)} \in \mathscr{K}(\mathscr{L}^{(2)})$; and so by the preceding, for any polynomial $P: R^1 \to R^1: P \circ \Phi_M \in \mathrm{Dom}(\mathscr{L})$ and $\mathscr{L}(P \circ \Phi_M) = 1/2\langle \Phi_M, \Phi_M \rangle_{\mathscr{L}} P'' \circ \Phi_M + (\mathscr{L}\Phi_M) P' \circ \Phi_M$. Thus if $F \in \mathscr{S}(R^1)$, and we approximate F in $C_b^2([-2M, 2M])$ by polynomials P_n, then we see that $P_n \circ \Phi_M \to F \circ \Phi_M$ and $\mathscr{L}(P_n \circ \Phi_M) \to 1/2\langle \Phi_M, \Phi_M \rangle_{\mathscr{L}} F'' \circ \Phi_M + (\mathscr{L}\Phi_M) P' \circ \Phi_M$ in $L^2(m)$ as $n \to \infty$. Hence $F \circ \Phi_M \in \mathrm{Dom}(\mathscr{L})$ and $\mathscr{L}(F \circ \Phi_M) = 1/2\langle \Phi_M, \Phi_M \rangle_{\mathscr{L}} F'' \circ \Phi_M + (\mathscr{L}\Phi_M) F' \circ \Phi_M$. Finally, if $\Phi = \Phi^{(1)} \otimes \Phi^{(2)}$, then $F \circ \Phi_M \to F \circ \Phi$, $F' \circ \Phi_M \to F' \circ \Phi$, and $F'' \circ \Phi_M \to F'' \circ \Phi$ boundedly as $M \nearrow \infty$. At the same time, using the results of the preceding paragraph, one sees that $\mathscr{L}\Phi_M \to \mathscr{L}\Phi$ and $\langle \Phi_M, \Phi_M \rangle_{\mathscr{L}} \to \langle \Phi, \Phi \rangle_{\mathscr{L}}$ in $L^2(m)$. Thus $F \circ \Phi_M \to F \circ \Phi$ in $L^2(m)$ and $\mathscr{L}(F \circ \Phi_M) \to 1/2\langle \Phi, \Phi \rangle_{\mathscr{L}} F'' \circ \Phi + (\mathscr{L}\Phi) F' \circ \Phi$ in $L^2(m)$. This proves that $F \circ \Phi \in \mathrm{Dom}(\mathscr{L})$ and that $\mathscr{L}(F \circ \Phi)$ is given by the desired expression. Q.E.D.

(2.2) Corollary. *For $n \geqslant 1$ let $(E^{(n)}, \mathscr{F}^{(n)}, m^{(n)})$ be a separable probability space and $(\mathscr{L}^{(n)}, T_\cdot^{(n)}, \mathscr{D}^{(n)}, m^{(n)})$ a symmetric diffusion semigroup on $L^2(m^{(n)})$. Set $E = \prod_1^\infty E^{(n)}$, $\mathscr{F} = \prod_1^\infty \mathscr{F}^{(n)}$, and $m = \prod_1^\infty m^{(n)}$; and define $T_\tau = \otimes_1^\infty T_\tau^{(n)}$ on $L^2(m)$. Then $\{T_\tau : \tau > 0\}$ is a symmetric Markov semigroup on $L^2(m)$. Moreover, if $\mathscr{L}$ is the generator of $\{T_\tau : \tau > 0\}$ and $\mathscr{D}$ is the span of the functions $\Phi^{(1)} \otimes \cdots \otimes \Phi^{(N)}$, $N \geqslant 1$ and $\Phi^{(n)} \in \mathscr{D}^{(n)}$ for $1 \leqslant n \leqslant N$, then $(\mathscr{L}, T_\cdot, \mathscr{D}, m)$ is a symmetric diffusion semigroup on $L^2(m)$. Finally, if $N \geqslant 1$ and $\Phi^{(n)} \in \mathrm{Dom}(\mathscr{L}^{(n)})$, $1 \leqslant n \leqslant N$, then $\Phi^{(1)} \otimes \cdots \otimes \Phi^{(N)} \in \mathrm{Dom}(\mathscr{L})$, $\mathscr{L}(\Phi^{(1)} \otimes \cdots \otimes \Phi^{(N)}) = \sum_1^N \Phi^{(1)} \otimes \cdots \otimes \Phi^{(n-1)} \otimes \mathscr{L}^{(n)}\Phi^{(n)} \otimes \Phi^{(n+1)} \otimes \cdots \otimes \Phi^{(N)}$, and $\langle \Phi^{(1)} \otimes \cdots \otimes \Phi^{(N)}, \Phi^{(1)} \otimes \cdots \otimes \Phi^{(N)} \rangle_{\mathscr{L}} = \sum_1^N (\Phi^{(1)})^2 \otimes \cdots \otimes (\Phi^{(n-1)})^2 \otimes \langle \Phi^{(n)}, \Phi^{(n)} \rangle_{\mathscr{L}^{(n)}} \otimes (\Phi^{(n+1)})^2 \otimes \cdots \otimes (\Phi^{(N)})^2$.*

(2.3) Theorem. *Let $(E, \mathscr{F}, m)$* and *$(\tilde{E}, \tilde{\mathscr{F}}, \tilde{m})$ be separable probability spaces and suppose that $\Xi: \tilde{E} \to E$ is a measurable measure preserving map with the property that the isometry $\Lambda: L^2(m) \to L^2(\tilde{m})$ given by $\Lambda\Phi = \Phi \circ \Xi$*

is onto. Given a symmetric diffusion semigroup $(\mathscr{L}, T_., \mathscr{D}, m)$ *on* $L^2(m)$, *define* $\tilde{\mathscr{L}} = \Lambda \circ \mathscr{L} \circ \Lambda^{-1}$ *on* $\Lambda(\mathrm{Dom}(\mathscr{L}))$, $\tilde{T}_\tau = \Lambda \circ T_\tau \circ \Lambda^{-1}$, *and* $\tilde{\mathscr{D}} = \Lambda\mathscr{D}$. *Then* $(\tilde{\mathscr{L}}, \tilde{T}_., \tilde{\mathscr{D}}, \tilde{m})$ *is a symmetric diffusion semigroup on* $L^2(\tilde{m})$.

Proof. Observe that if $\Phi = (\Phi_1,..., \Phi_n) \in (L^2(m))^n$ and if $F: R^n \to R$ is a bounded measurable function, then $\Lambda(F \circ \Phi) = F \circ (\Lambda\Phi)$, where $\Lambda\Phi = (\Lambda\Phi_1,..., \Lambda\Phi_n)$. Hence, if $\tilde{\Phi} = (\tilde{\Phi}_1,..., \Phi_n) \in L^2(\tilde{m})$, then $\Lambda^{-1}(F \circ \tilde{\Phi}) = F \circ (\Lambda^{-1}\tilde{\Phi})$. From these facts it is completely elementary to verify that $(\tilde{\mathscr{L}}, \tilde{T}_., \tilde{\mathscr{D}}, \tilde{m})$ is a symmetric diffusion semigroup. Q.E.D.

Let $\gamma(dx) = (1/(2\pi)^{1/2})\, e^{-x^2/2}\, dx$ on R^1 and define

$$T_\tau \Phi(x) = \int p(1 - e^{-\tau}, y - xe^{-\tau/2})\, \Phi(y)\, dy, \quad \tau > 0 \qquad \text{and} \qquad \Phi \in C_b(R^1),$$

where $p(s, \xi) = (1/(2\pi)^{1/2})\, e^{-\xi^2/2s}$. Then it is easy to check that $\{T_\tau : \tau > 0\}$ is a semigroup and that for each $\tau > 0$:

$$\int \Psi T_\tau \Phi\, d\gamma = \int \Phi T_\tau \Psi\, d\gamma, \qquad \Phi, \Psi \in C_b(R^1).$$

Since T_τ is non-negativity preserving and $T_\tau 1 = 1$, it follows that each T_τ has a unique extention T_τ^γ as a contraction on $L^2(\gamma)$ and that the resulting family of extentions $\{T_\tau^\gamma : \tau > 0\}$ forms a symmetric Markov semigroup. Let $\mathscr{L}^\gamma$ denote the generator of $\{T_\tau^\gamma : \tau > 0\}$.

For $n \geqslant 0$, define

$$H_n(x) = \frac{(-1)^n}{(n!)^{1/2}}\, e^{x^2/2} D^n(e^{-x^2/2}). \tag{2.4}$$

Using the generating function

$$e^{\lambda x - \lambda^2/2} = \sum_0^\infty \frac{\lambda^n}{(n!)^{1/2}} H_n(x), \qquad \lambda \in R^1 \qquad \text{and} \qquad x \in R^1, \tag{2.5}$$

one easily derives:

$$T_\tau^\gamma H_n = e^{-n\tau/2} H_n, \qquad n \geqslant 0 \qquad \text{and} \qquad \tau > 0. \tag{2.6}$$

Since $\{H_n : n \geqslant 0\}$ is an orthonormal basis in $L^2(\gamma)$, it follows from (2.6) that:

$$T_\tau^\gamma \Phi = \sum_0^\infty e^{-n\tau/2} (\Phi, H_n)_{L^2(\gamma)} H_n, \qquad \Phi \in L^2(\gamma), \tag{2.7}$$

and

$$\mathscr{L}^\gamma \Phi = -\sum_0^\infty \frac{n}{2} (\Phi, H_n)_{L^2(\gamma)} H_n, \qquad \Phi \in \mathrm{Dom}(\mathscr{L}^\gamma), \tag{2.8}$$

where

$$\mathrm{Dom}(\mathscr{L}^\gamma) \equiv \left\{ \Phi \in L^2(\gamma) : \sum_0^\infty n^2 (\Phi, H_n)^2_{L^2(\gamma)} < \infty \right\}. \tag{2.9}$$

(2.10) LEMMA. *$C^2_\nearrow(R^1) \subseteq \mathrm{Dom}(\mathscr{L}^\gamma)$ and for $\Phi \in C^2_\nearrow(R^1)$:*

$$\mathscr{L}^\gamma \Phi = 1/2(D^2 - xD)\,\Phi. \tag{2.11}$$

Proof. First recall that if $h_n(x) = H_n(x)\, e^{-x^2/4}$, then $f \in \mathscr{S}(R^1)$ if and only $\{(f, h_n)_{L^2(R^1)} : n \geqslant 0\}$ is rapidly decreasing. Moreover, if $f \in \mathscr{S}(R^1)$, then $\sum_0^N (f, h_n)_{L^2(R^1)}\, h_n \to f$ in $\mathscr{S}(R^1)$. Thus $e^{x^2/4}\mathscr{S}(R^1) = \{f \in L^2(\gamma) : (f, H_n)_{L^2(\gamma)}$ is rapidly decreasing$\}$, and for $f \in e^{x^2/4}\mathscr{S}(R^1)$: $e^{-e^2/4} \sum_0^N (f, H_n)_{L^2(\gamma)}\, H_n \to e^{-x^2/4} f$ in $\mathscr{S}(R^1)$. Since $C^\infty_\nearrow(R^1) \subseteq e^{x^2/4}\mathscr{S}(R^1)$, we now see that for $f \in C^\infty_\nearrow(R^1)$: $f \in D(\mathscr{L})$ and

$$\begin{aligned}
1/2(D^2 - xD) f &= \lim_{N \nearrow \infty} \sum_0^N (f, H_n)_{L^2(\gamma)} (1/2(D^2 - xD))\, H_n \\
&= \lim_{N \nearrow \infty} -\sum_0^N \frac{n}{2} (f, H_n)_{L^2(\gamma)}\, H_n = \mathscr{L}^\gamma f.
\end{aligned}$$

To complete the proof, simply observe that if $f \in C^2_\nearrow(R^1)$, then there is a sequence $\{f_n\}_1^\infty \subseteq C^\infty_\nearrow(R^1)$ such that the pairs $(f_n, 1/2(D^2 - xD) f_n) \to (f, 1/2(D^2 - xD) f)$ in $(L^2(\gamma))^2$. Q.E.D.

(2.12) THEOREM. *Let $\mathscr{D}^\gamma$ be the set of real polynomials on R^1. Then $(\mathscr{L}^\gamma, T^\gamma_\cdot, \mathscr{D}^\gamma, \gamma)$ is a symmetric diffusion semigroup on $L^2(\gamma)$.*

Proof. Clearly all that we need to do is check that if $F \in \mathscr{S}(R^1)$ and $\Phi \in \mathscr{D}^\gamma$, then $F \circ \Phi \in D(\mathscr{L}^\gamma)$ and $\mathscr{L}^\gamma (F \circ \Phi) = 1/2\langle \Phi, \Phi \rangle_{\mathscr{L}^\gamma} F'' \circ \Phi + (\mathscr{L}^\gamma \Phi)\, F' \circ \Phi$. But, by Lemma (2.10):

$$\begin{aligned}
\langle \Phi, \Phi \rangle_{\mathscr{L}^\gamma} &= (\Phi')^2, \\
\mathscr{L}^\gamma \Phi &= 1/2(\Phi'' - x\Phi'),
\end{aligned}$$

and

$$\mathscr{L}^\gamma (F \circ \Phi) = 1/2(\Phi')^2\, F'' \circ \Phi + 1/2(\Phi'' - x\Phi')\, F' \circ \Phi. \qquad \text{Q.E.D.}$$

(2.13) COROLLARY. *Let $\Gamma = \gamma^{Z^+}$ on $(R^{Z^+}, \mathscr{B}_{R^{Z^+}})$ and $T^\Gamma_\tau = (T^\gamma_\tau)^{Z^+}$, $\tau > 0$, on $L^2(\Gamma)$. Then $\{T^\Gamma_\tau : \tau > 0\}$ is a symmetric Markov semigroup on $L^2(\Gamma)$. Moreover, if $\mathscr{L}^\Gamma$ is the generator of $\{T^\Gamma_\tau : \tau > 0\}$ and $\mathscr{D}^\Gamma$ denotes the algbra of real polynomials on R^{Z^+}, then $(\mathscr{L}^\Gamma, T^\Gamma_\cdot, \mathscr{D}^\Gamma, \Gamma)$ is a symmetric diffusion* *semigroup on* $L^2(\Gamma)$. *Finally, if* $N \geqslant 1, f \in C^2_{\nearrow}(R^N)$, *and* $\Phi(x) = f(x_1,\dots,x_N)$, $x \in R^{Z^+}$, *then* $\Phi \in \mathscr{K}(\mathscr{L}^\Gamma)$,

$$\mathscr{L}^\Gamma \Phi = 1/2 \sum_{n=1}^{N} \left(\frac{\partial^2 f}{\partial x_n^2} - x_n \frac{\partial f}{\partial x_n}\right)(x_1,\dots,x_N),$$

and $\langle \Phi, \Phi \rangle_{\mathscr{L}^\Gamma} = \sum_1^N (\partial f/\partial x_n)^2 (x_1,\dots,x_N)$.

It is easy to provide a complete description of the spectrum of $\mathscr{L}^\Gamma$. Indeed, let $A = \{\alpha \in \mathscr{N}^{Z^+}: |\alpha| = \sum_1^\infty |\alpha_k| < \infty\}$, where $\mathscr{N} = \{0, 1,\dots, n,\dots\}$, and for $\alpha \in A$ define

$$H_\alpha(x) = \prod_1^\infty H_{\alpha_k}(x_k), \quad x \in R^{Z^+}, \tag{2.14}$$

where $\prod_1^\infty H_{\alpha_k}(x) \equiv \prod_{n \in [\alpha]} H_{\alpha_k}(x_n)$ with $[\alpha] \equiv \{k \in Z^+ : \alpha_k > 0\}$. Clearly $H_\alpha \in D(\mathscr{L}^\Gamma)$ and

$$\mathscr{L}^\Gamma H_\alpha = \frac{-|\alpha|}{2} H_\alpha$$

for each $\alpha \in A$. Moreover, if

$$\mathscr{H}^{(n)} = \overline{\operatorname{span}\{H_\alpha : \alpha \in A \text{ and } |\alpha| = n\}}^{L^2(\Gamma)}, \tag{2.15}$$

then the $\mathscr{H}^{(n)}$'s are mutually orthogonal subspaces of $L^2(\Gamma)$ and $L^2(\Gamma) = \bigoplus_0^\infty \mathscr{H}^{(n)}$. Hence, if $E_{\mathscr{H}^{(n)}}$ denotes the orthogonal projection operator in $L^2(\Gamma)$ onto $\mathscr{H}^{(n)}$, then

$$\operatorname{Dom}(\mathscr{L}^\Gamma) = \left\{\Phi \in L^2(\Gamma): \sum_0^\infty n^2 \|E_{\mathscr{H}^{(n)}} \Phi\|^2_{L^2(\Gamma)} < \infty\right\}$$

and

$$\mathscr{L}^\Gamma \Phi = -\sum_0^\infty \frac{n}{2} E_{\mathscr{H}^{(n)}} \Phi, \qquad \Phi \in \operatorname{Dom}(\mathscr{L}^\Gamma).$$

(2.16) THEOREM. *Let* $U = ((u_{k,l}))_{k,l \in Z^+}$ *be a real orthogonal matrix on* $l^2(Z^+)$. *Then* $(Ux)_k = \sum_{l \in Z^+} u_{k,l} x_l$ *converges* (*a.s.*, Γ) *for all* $k \in Z$. *Moreover, if* $U: R^{Z^+} \to R^{Z^+}$ *is defined by* $(Ux)_k = \sum_{l \in Z^+} u_{k,l} x_l$ *for those* $x \in R^{Z^+}$ *such that* $\sum u_{k,l} x_l$ *converges and* $(Ux)_k = 0$ *otherwise, then* U *is measure preserving. Finally, if* $\Lambda_U: L^2(\Gamma) \to L^2(\Gamma)$ *is the isometry given by* $\Lambda_U \Phi = \Phi \circ U$, *then* $\Lambda_U \circ \Lambda_{U^*} = \Lambda_{U^*} \circ \Lambda_U =$ *identity and* $\Lambda_U \mathscr{H}^{(n)} = \mathscr{H}^{(n)}$ *for all* $n \geqslant 0$. *In particular,* $\mathscr{L}^\Gamma \circ \Lambda_U = \Lambda_U \circ \mathscr{L}^\Gamma$ *and* $T^\Gamma_\tau \circ \Lambda_U = \Lambda^U \circ T^\Gamma_\tau$, $\tau > 0$.

Proof. The convergence assertion is an immediate consequence of Kolmogorov's three series theorem. Furthermore, since $\{x_k : k \in Z^+\}$ is a Gaussian family on $(R^{Z^+}, \mathscr{B}_{R^{Z^+}}, \Gamma)$ and $E^\Gamma[(Ux)_k(Ux)_l] = \delta_{k,l}$, it follows that U is measure preserving. To prove that $\Lambda_{U^*} \circ \Lambda_U = \text{identity}$, suppose $\Phi(x) = f(x_1, \ldots, x_N)$ for some $N \geqslant 1$ and $f \in C_b(R^N)$, and for $L \geqslant N$ define $\Phi_L(x) = f(\sum_{l1}^L u_{1l}x_l, \ldots, \sum_{l=1}^L u_{Nl}x_l)$. Then $\Phi_L \to \Lambda_U \Phi$ in $L^2(\Gamma)$ as $L \nearrow \infty$. At the same time, $\sum_{k=1}^\infty u_{k,m} \sum_{l=1}^L u_{k,l}x_l = x_m$ for $1 \leqslant m \leqslant N \leqslant L$. Thus $\Lambda_{U^*}\Phi_L = \Phi$ for all $L \geqslant N$ and so $\Lambda_{U^*} \circ \Lambda_U \Phi = \lim_{L \nearrow \infty} \Lambda_{U^*}\Phi_L = \Phi$. Since the set of such Φ's is dense in $L^2(\Gamma)$, we now see that $\Lambda_{U^*} \circ \Lambda_U = \text{identity}$. Clearly this also implies that $\Lambda_U \circ \Lambda_{U^*} = \text{identity}$. Thus, in view of the preceding discussion about the spectrum of $\mathscr{L}^\Gamma$, the proof will be complete once we show that $\Lambda_U \mathscr{H}^{(n)} \subseteq \mathscr{H}^{(n)}$ for all $n \geqslant 0$. In particular, what we must check is that if $n \geqslant 0$ and $v \in R^{Z^+}$ satisfies $|v| = 1$, then the function $\Phi(x) = H_n(\sum_1^\infty v_k x_k)$ is in $\mathscr{H}^{(n)}$. By a simple approximation argument it is enough for us to do this in the case when $v_k = 0$ for all k greater than some L. But using the generating function (2.5), one easily derives: $H_n(\sum_1^L v_k x_k) = \sum_{|\alpha| = n} \binom{n}{\alpha}^{1/2} v^\alpha H_\alpha(x)$ for any $v \in R^L$ with $|v| = 1$. In particular, $H_n(\sum_1^L v_k x_k) \in \mathscr{H}^{(n)}$. Thus $\Lambda_U H_\alpha \in \bigoplus_0^{|\alpha|} \mathscr{H}^{(m)}$. At the same time, if $|\beta| < |\alpha|$, then $(\Lambda_U H_\alpha, H_\beta)_{L^2(\Gamma)} = (H_\alpha, \Lambda_U H_\beta)_{L^2(\Gamma)} = 0$, since $\Lambda_U H_\beta \in \bigoplus_0^{|\beta|} \mathscr{H}^{(m)}$. We therefore see that $\Lambda H_\alpha \in \mathscr{H}^{(|\alpha|)}$. Q.E.D.

3. The Ornstein–Uhlenbeck Semigroup on Wiener Space

Let Θ be the space of continuous functions θ: $[0, \infty) \to R^1$ satisfying $\theta(0) = 0$, and think of Θ as a Polish space with the metric of uniform convergence on compacts. Denote by $\mathscr{B}$ the Borel field over Θ and for each $t \geqslant 0$ let $\mathscr{B}_t$ denote the sub σ-algebra $\sigma(\theta(s): 0 \leqslant s \leqslant t)$. Finally, let $\mathscr{W}$ denote the Wiener measure on $(\Theta, \mathscr{B})$. That is, $\mathscr{W}$ is the unique Gaussian measure on $(\Theta, \mathscr{B})$ such that $E^{\mathscr{W}}[\theta(t)] = 0$ for all $t \geqslant 0$ and $E^{\mathscr{W}}[\theta(s)\,\theta(t)] = s \wedge t$ for all $s, t \geqslant 0$.

Given $n \geqslant 1$ and $t > 0$, set $\Delta_n(t) = \{(t_1, \ldots, t_n) \in R^n : 0 \leqslant t_1 \leqslant \cdots \leqslant t_n \leqslant t\}$ and $\Delta_n = \bigcup_{t \geqslant 0} \Delta_n(t)$. Given $f_1, \ldots, f_n \in L^2([0, \infty))$, define

$$\int_{\Delta_n(t)} f_1(t_1) \cdots f_n(t_n)\, d^n\theta$$

inductively by

$$\int_{\Delta_{n+1}(t)} f_1(t_1) \cdots f_{n+1}(t_{n+1})\, d^{n+1}\theta = \int_0^t f_{n+1}(t_{n+1}) \left(\int_{\Delta_n(t_{n+1})} f_1(t_1) \cdots f_n(t_n)\, d^n\theta \right) d_\theta(t_{n+1}).$$

Here and throughout, $d\theta(s)$-integrals are taken in the sense of Itô. For f's of the form $\sum_{l=1}^{L} f_{1,l}(t_1) \cdots f_{n,l}(t_n)$ with

$$\{f_{j,l}\}_{\substack{1 \leqslant j \leqslant n \\ 1 \leqslant l \leqslant L}} \subseteq L^2([0, \infty)),$$

we define $\int_{\Delta_n(t)} f\, d^n\theta$ by linearity. Using induction and elementary properties of Itô integrals, one finds that

$$\begin{aligned} &\text{(i)} \quad (\textstyle\int_{\Delta_n(t)} f\, d^n\theta, \mathscr{B}_t, \mathscr{W}) \text{ is a continuous square integrable} \\ &\qquad \text{martingale with mean 0,} \\ &\text{(ii)} \quad (\textstyle\int_{\Delta_m(s)} f\, d^m\theta, \int_{\Delta_n(t)} g\, d^n\theta)_{L^2(\mathscr{W})} = \delta_{m,n}(f, g)_{L^2(\Delta_m(s \wedge t))}, \end{aligned} \tag{3.1}$$

for f and g of the form just described. In particular, (3.1)(ii) shows that for each $t \geqslant 0$ and $n \geqslant 1$ there is a unique isometry $f \to \int_{\Delta_n(t)} f\, d^n\theta$ from $L^2(\Delta_n(t))$ into $L^2(\mathscr{W})$ such that $\int_{\Delta_n(t)} f\, d^n\theta$ is given by $\int_{\Delta_n(t)} f_1(t_1) \cdots f_n(t_n)\, d^n\theta$ when $f = f_1(t_1) \cdots f_n(t_n)$ with $\{f_j\}_1^n \subseteq L^2([0, \infty))$. Furthermore, it is easily seen that the properties in (3.1) continue to hold for all $f \in L^2(\Delta_m)$ and $g \in L^2(\Delta_m)$. Noting that $\| \int_{\Delta_n(t)} f\, d^n\theta \|_{L^2(\mathscr{W})} \leqslant \| f \|_{L^2(\Delta_n)}$, we see from the martingale convergence theorem that $\int_{\Delta_n} f\, d^n\theta \equiv \lim_{t \nearrow \infty} \int_{\Delta_n(t)} f\, d^n\theta$ exists, where the limit is taken in the sense of (a.s., $\mathscr{W}$) on $L^2(\mathscr{W})$ convergence. Moreover, $\int_{\Delta_n(t)} f\, d^n\theta = E^{\mathscr{W}}[\int_{\Delta_n} f\, d^n\theta \mid \mathscr{B}_t]$ (a.s., $\mathscr{W}$) for all $t \geqslant 0$.

Next, for $t \geqslant 0$, define $Z^{(0)}(t)$ to be the constants and $Z^{(n)}(t) = \{\int_{\Delta_n(t)} f\, d^n\theta : f \in L^2(\Delta_n)\}$, $n \geqslant 1$. For fixed $t \geqslant 0$, the $Z^{(n)}(t)$, $n \geqslant 0$, are mutually orthogonal closed linear subspaces of $L^2(\Theta, \mathscr{B}_t, \mathscr{W})$. Also, for fixed $n \geqslant 0$, the $Z^{(n)}(t)$, $t \geqslant 0$, are non-decreasing subspaces of $L^2(\mathscr{W})$ and $Z^{(n)} \equiv \overline{\bigcup_{t \geqslant 0} Z^{(n)}(t)}^{L^2(\mathscr{W})}$ coincides with $\{\int_{\Delta_n} f\, d^n\theta : f \in L^2(\Delta_n)\}$. Clearly the $Z^{(n)}$'s are mutually orthogonal in $L^2(\mathscr{W})$.

(3.2) Lemma (Wiener). *For each* $t \geqslant 0$, $L^2(\Theta, \mathscr{B}_t, \mathscr{W}) = \oplus_0^\infty Z^{(n)}(t)$. *In particular,* $L^2(\mathscr{W}) = \oplus_0^\infty Z^{(n)}$.

Proof. Let $f: [0, \infty) \to R^1$ be bounded and measurable. Then for $\lambda \in \mathbb{C}$,

$$\begin{aligned} X_\lambda(t) &= \left[\lambda \int_0^t f(s)\, d\theta(s) - \lambda^2/2 \int_0^t f^2(s)\, ds \right] \\ &= 1 + \sum_1^\infty \lambda^n \int_0^t f^{(n)}\, d^{(n)}\theta, \qquad t \geqslant 0, \end{aligned} \tag{3.3}$$

where $f^{(n)} = f(t_1) \cdots f(t_n)$ and the convergence of the series on the right is in L^2 uniformly for t in compacts. The proof of (3.3) goes as follows. By Itô's formula:

$$\| X_\lambda(t) \|_{L^2(\mathscr{W})}^2 \leqslant \exp\left((\operatorname{Re} \lambda)^2 \int_0^t f^2(s)\, ds \right)$$

and

$$X_\lambda(t) = 1 + \lambda \int_0^t f(s)\, X_\lambda(s)\, d\theta(s), \qquad t \geq 0. \tag{3.4}$$

Since (3.4) admits only one solution, (3.3) is now proved.

Now suppose that $\Phi \in L^2(\Theta, \mathscr{B}_t, \mathscr{W})$ and that $\Phi \perp \oplus_0^\infty Z^{(n)}(t)$. Then by (3.3):

$$\left(\Phi, \exp\left(\lambda \int_0^t f(s)\, d\theta - \lambda^2/2 \int_0^t f^2(s)\, ds\right)\right)_{L^2(\mathscr{W})} = 0$$

for all bounded measurable $f\colon [0, \infty) \to R^1$. But this means that

$$\left(\Phi, \exp\left(i \sum_1^n \alpha_j(\theta(t_j) - \theta(t_{j-1}))\right)\right)_{L^2(\mathscr{W})} = 0$$

for all $n \geq 1$, $0 \leq t_0 < \cdots < t_n \leq t$, and $\alpha_1, ..., \alpha_n \in R^1$. Hence $\Phi \equiv 0$. Q.E.D.

We now define the Ornstein–Uhlenbeck semigroup on Wiener space. Namely, let $E_{Z^{(n)}}$ denote orthogonal projection in $L^2(\mathscr{W})$ onto $Z^{(n)}$ and set

$$D(\mathscr{L}^{\mathscr{W}}) = \left\{ \Phi \in L^2(\mathscr{W}) \colon \sum_0^\infty n^2 \, \|E_{Z^{(n)}} \Phi\|^2_{L^2(\mathscr{W})} < \infty \right\}, \tag{3.5}$$

and

$$\mathscr{L}^{\mathscr{W}} \Phi = -\sum_0^\infty n/2 E_{Z^{(n)}} \Phi, \qquad \Phi \in D(\mathscr{L}^{\mathscr{W}}). \tag{3.6}$$

Clearly $\mathscr{L}^{\mathscr{W}}$ is the generator of the $L^2(\mathscr{W})$-semigroup

$$T_\tau^{\mathscr{W}} \Phi = \sum_0^\infty e^{-n\tau/2} E_{Z^{(n)}} \Phi, \qquad \tau > 0 \quad \text{and} \quad \Phi \in L^2(\mathscr{W}), \tag{3.7}$$

and obviously the $T_\tau^{\mathscr{W}}$'s are the self-adjoint contractions. It is less obvious that $\mathscr{L}^{\mathscr{W}}$ is the generator of a symmetric diffusion semigroup.

(3.8) Lemma. *Let $\mathscr{F} = \{f_n\}_1^\infty$ be a real orthonormal basis in $L^2([0, \infty))$ and define a map $\mathscr{F}\colon \Theta \to R^{Z^+}$ so that $(\mathscr{F}(\theta))_k = \int_0^\infty f_k(s)\, d\theta(s)$, $k \in Z^+$. Then $\mathscr{F}$ is measure preserving from $(\Theta, \mathscr{B}, \mathscr{W})$ into $(R^{Z^+}, \mathscr{B}, \Gamma)$. Moreover, the isometry $A_{\mathscr{F}}\colon L^2(\Gamma) \to L^2(\mathscr{W})$ given by $A_{\mathscr{F}}\Psi = \Psi \circ \mathscr{F}$ is onto. Finally, if $\mathscr{G} = \{g_n\}_1^\infty$ is a second such basis, then $A_{\mathscr{F}}(\mathscr{H}^{(n)}) = A_{\mathscr{G}}(\mathscr{H}^{(n)})$ for all $n \geq 0$.*

Proof. To see that $\mathscr{F}$ is measure preserving simply note that $\{\int_0^\infty f_k\, d\theta \colon k \geq 1\}$ is a mean 0 Gaussian family on $(\Theta, \mathscr{B}, \mathscr{W})$ and that $E^{\mathscr{W}}[\int_0^\infty f_k\, d\theta \int_0^\infty f_l\, d\theta] = \delta_{kl}$.

We next prove the final assertion of the lemma. Define

$U = (((f_k, g_l)))_{k,l \in Z^+}$. By Theorem (2.16), $\Lambda_U : L^2(\Gamma) \to L^2(\Gamma)$ is a unitary map and $\Lambda_U(\mathscr{H}^{(n)}) = \mathscr{H}^{(n)}$, $n \geqslant 0$. Hence, since $\Lambda_{\mathscr{F}} = \Lambda_{\mathscr{G}} \circ \Lambda_U$, it is clear that $\Lambda_{\mathscr{F}}(\mathscr{H}^{(n)}) = \Lambda_{\mathscr{G}}(\mathscr{H}^{(n)})$, $n \geqslant 0$.

Finally, to see that $\Lambda_{\mathscr{F}}(L^2(\Gamma)) = L^2(\mathscr{W})$, note first that, by the preceding, $\Lambda_{\mathscr{F}}(L^2(\Gamma))$ is independent of the choice of basis $\mathscr{F}$. In particular, if $0 \leqslant t_0 < t_1 < \cdots < t_n$ and $\mathscr{F}$ is chosen so that $f_m = (t_m - t_{m-1})^{-1/2} \chi_{[t_{m-1}, t_m)}$, $1 \leqslant m \leqslant n$, then for any $F \in C_b(R^n)$:

$$F(\theta(t_1) - \theta(t_0), \ldots, \theta(t_n) - \theta(t_{n-1})) = \Lambda_{\mathscr{F}} \tilde{F},$$

where $\tilde{F}(x) = F((t_1 - t_0)^{1/2} x_1, \ldots, (t_n - t_{n-1})^{1/2} x_n)$. Q.E.D.

(3.9) THEOREM. *Let $\mathscr{F} = \{f_n\}_1^\infty$ be a real orthonormal basis in $L^2([0, \infty))$ and define $\Lambda_{\mathscr{F}}: L^2(\Gamma) \to L^2(\mathscr{W})$ accordingly (as in Lemma (3.8)). Then $T_\tau^{\mathscr{W}} = \Lambda_{\mathscr{F}} \circ T_\tau^\Gamma \circ \Lambda_{\mathscr{F}}^{-1}$, $\tau > 0$, and $\mathscr{L}^{\mathscr{W}} = \Lambda_{\mathscr{F}} \circ \mathscr{L}^\Gamma \circ \Lambda_{\mathscr{F}}^{-1}$, where $T_\tau^{\mathscr{W}}$ and $\mathscr{L}^{\mathscr{W}}$ are given by (3.7) and (3.6), respectively. In particular, $(\mathscr{L}^{\mathscr{W}}, T_\cdot^{\mathscr{W}}, \Lambda_{\mathscr{F}} \mathscr{D}^\Gamma, \mathscr{W})$ is a symmetric diffusion semigroup on $L^2(\mathscr{W})$. Finally, if $t \geqslant 0$, Φ is a $\mathscr{B}_t$-measurable element of $L^2(\mathscr{W})$, and Ψ is a $\mathscr{B}^t \equiv \sigma(\theta(s) - \theta(t): s \geqslant t)$-measurable element of $L^2(\mathscr{W})$, then $\Phi \cdot \Psi \in L^2(\mathscr{W})$, $T_\tau^{\mathscr{W}} \Phi$ is $\mathscr{B}_t$-measurable, $T_\tau^{\mathscr{W}} \Psi$ is $\mathscr{B}^t$-measurable, and $T_\tau^{\mathscr{W}}(\Phi \cdot \Psi) = (T_\tau^{\mathscr{W}} \Phi) \cdot (T_\tau^{\mathscr{W}} \Psi)$. In particular, if $\Phi, \Psi \in \mathrm{Dom}(\mathscr{L}^{\mathscr{W}})$, Φ is $\mathscr{B}_t$-measurable and Ψ is $\mathscr{B}^t$-measurable, then $\Phi \cdot \Psi \in \mathrm{Dom}(\mathscr{L}^{\mathscr{W}})$ and $\mathscr{L}^{\mathscr{W}}(\Phi \cdot \Psi) = \Phi \mathscr{L}^{\mathscr{W}} \Psi + \Psi \mathscr{L}^{\mathscr{W}} \Phi$.*

Proof. We first check that $T_\tau \equiv \Lambda_{\mathscr{F}} \circ T_\tau^\Gamma \circ \Lambda_{\mathscr{F}}^{-1}$ is independent of the choice of basis $\mathscr{F}$. To this end, let E_n denote orthogonal projection in $L^2(\mathscr{W})$ onto $\Lambda_{\mathscr{F}}(\mathscr{H}^{(n)})$. By Lemma (3.8), E_n does not depend on $\mathscr{F}$; and by the spectral properties of T_τ^Γ, $\Lambda_{\mathscr{F}} \circ T_\tau^\Gamma \circ \Lambda_{\mathscr{F}}^{-1} = \sum_0^\infty e^{-n\tau/2} E_n$. Thus T_τ is indeed defined independent of the choice of $\mathscr{F}$.

We next prove the final assertion of the theorem with T_τ replacing $T_\tau^{\mathscr{W}}$. Given $t > 0$, choose $\mathscr{F}$ so that $\mathrm{supp}(f_{2n+1}) \subseteq [0, t]$ and $\mathrm{supp}(f_{2n}) \subseteq [t, \infty)$ for all $n \geqslant 1$ and; define $\mathscr{H}^+$ and $\mathscr{H}^-$ to be, respectively, the subspaces of $L^2(\Gamma)$ consisting of $F \in L^2(\Gamma)$ such that $F(x) = F(x_1, x_3, \ldots, x_{2n+1}, \ldots)$ and $F(x) = F(x_2, x_4, \ldots, x_{2n}, \ldots)$. Then, by Lemma (3.2), $\Lambda^{\mathscr{F}}(\mathscr{H}^+) = L^2(\Theta, \mathscr{B}_t, \mathscr{W})$ and $\Lambda_{\mathscr{F}}(\mathscr{H}^-) = L^2(\Theta, \mathscr{B}^t, \mathscr{W})$. Moreover, it is easy to see from the construction of T_τ^Γ that $\mathscr{H}^+$ and $\mathscr{H}^-$ are invariant under T_τ^Γ and that $T_\tau^\Gamma(\Psi^+ \cdot \Psi^-) = (T_\tau^\Gamma \Psi^+) \cdot (T_\tau^\Gamma \Psi^-)$ for $\Psi^\pm \in \mathscr{H}^\pm$. Hence the desired properties of T_τ can now be read off from the representation $T_\tau = \Lambda_{\mathscr{F}} \circ T_\tau^\Gamma \circ \Lambda_{\mathscr{F}}^{-1}$.

To complete the proof we must still identify the generator $\mathscr{L}$ of $\{T_\tau : \tau > 0\}$ as $\mathscr{L}^{\mathscr{W}}$; and, in view of the equality $L^2(\mathscr{W}) = \bigoplus_0^\infty Z^{(n)}$, this identification will have been made once we show that $Z^{(n)} \subseteq \mathrm{Dom}(\mathscr{L})$ and $\mathscr{L}\Phi = -(n/2)\,\Phi$, $\Phi \in Z^{(n)}$, for all $n \geqslant 0$. Clearly there is nothing to do when $n = 0$. In order to handle $n \geqslant 1$, we first show that

$\mathscr{L}(\theta(t+h)-\theta(t))=-1/2(\theta(t+h)-\theta(t))$ for any $t\geqslant 0$ and $h>0$. To this end, choose $\mathscr{F}$ so that $f_1=h^{-1/2}\chi_{[t,t+h)}$. Then $\mathscr{L}(\theta(t+h)-\theta(t))=h^{1/2}\mathscr{L}(\int_0^\infty f_1\,d\theta)=h^{1/2}\Lambda_{\mathscr{F}}^{-1}(\mathscr{L}^\Gamma x_1)=-1/2h^{1/2}\Lambda_{\mathscr{F}}^{-1}(x_1)=-1/2(\theta(t+h)-\theta(t))$. We now proceed by induction on n. That is, assume that $Z^{(n)}\subseteq \operatorname{Dom}(\mathscr{L})$ and $\mathscr{L}\Phi=-n/2\Phi$, $\Phi\in Z^{(n)}$. Given $g_1,\dots,g_{n+1}\in C_0([0,\infty))$, set $\Phi(t)=\int_{\Delta_n(t)}g_1\cdots g_n d^{(n)}\theta$. By induction hypothesis, $\Phi(t)\in\operatorname{Dom}(\mathscr{L})$ and $\mathscr{L}\Phi(t)=-(n/2)\,\Phi(t)$ for all $t\geqslant 0$. Moreover, by the preceding paragraph, $\Phi(t)\,g_{n+1}(t)(\theta(t+h)-\theta(t))\in\operatorname{Dom}(\mathscr{L})$ and $\mathscr{L}(\Phi(t)\,g_{n+1}(t)(\theta(t+h)-\theta(t)))=\mathscr{L}(\Phi(t))(g_{n+1}(t)(\theta(t+h)-\theta(t)))+\Phi(t)\,g_{n+1}(t)\,\mathscr{L}(\theta(t+h)-\theta(t))$ for all $t\geqslant 0$ and $h>0$. Combining this with the observation just made, we now see that $\mathscr{L}(\Phi(t)\,g_{n+1}(t)(\theta(t+h)-\theta(t)))=-((n+1)/2)\,\Phi(t)\,g_{n+1}(t)(\theta(t+h)-\theta(t))$. Hence for any $N\geqslant 1$:

$$\Phi_N\equiv\sum_1^{N^2}\Phi(k/N)\,g_{n+1}(k/N)\left(\theta\left(\frac{k+1}{N}\right)-\theta\left(\frac{k}{N}\right)\right)\in\operatorname{Dom}(\mathscr{L})$$

and $\mathscr{L}\Phi_N=-((n+1)/2)\,\Phi_N$. Since $\Phi_N\to\int_{\Delta_{n+1}}g_1\cdots g_{n+1}d^{n+1}\theta$ in $L^2(\mathscr{W})$ as $N\to\infty$, we conclude that $\int_{\Delta_{n+1}}g_1\cdots g_{n+1}d^{n+1}\theta\in\operatorname{Dom}(\mathscr{L})$ and that $\mathscr{L}(\int_{\Delta_{n+1}}g_1\cdots g_{n+1}d^{n+1}\theta)=-((n+1)/2)\int_{\Delta_{n+1}}g_1\cdots g_{n+1}d^{n+1}\theta$. From here it is an easy matter to check that $Z^{(n+1)}\subseteq\operatorname{Dom}(\mathscr{L})$ and that $\mathscr{L}\Phi=-((n+1)/2)\,\Phi$, $\Phi\in Z^{(n+1)}$. Q.E.D.

(3.10) *Remark.* It may not be entirely obvious why the preceding semigroup should be called the Ornstein–Uhlenbeck semigroup on Wiener space. Indeed, given a Gaussian measure μ on R^d with mean 0 and covariance A, the associated Ornstein–Uhlenbeck operator is

$$1/2\left(\sum_{i,j=1}^{d}A_{ij}\frac{\partial^2}{\partial x_i\,\partial x_j}-\sum_1^d x_i\frac{\partial}{\partial x_i}\right).$$

Thus the Ornstein–Uhlenbeck operator on Wiener space should be

$$1/2\left(\iint s\wedge t\frac{\partial^2}{\partial\theta(s)\,\partial\theta(t)}\,ds\,dt-\int\theta(s)\frac{\partial}{\partial\theta(s)}\,ds\right). \tag{3.11}$$

It turns out that the operator $\mathscr{L}^{\mathscr{W}}$ is in fact an extension of (3.11). Details of this appraoch to $\mathscr{L}^{\mathscr{W}}$ can be found in Shigekawa [11]. In order to see the relation between $\mathscr{L}^{\mathscr{W}}$ and the operator in (3.11), we will content ourselves with the following example. Let $0\leqslant t_1<\cdots<t_n$ and $f\in C^2_{\not\uparrow}(R^n)$ be given and set $\Phi=f(\theta(t_1),\dots,\theta(t_n))$. Then $\Phi\in\operatorname{Dom}(\mathscr{L}^{\mathscr{W}})$ and

$$\mathscr{L}^{\mathscr{W}}=1/2\sum_{i,j=1}^{n}\langle\theta(t_i),\theta(t_j)\rangle_{\mathscr{L}^{\mathscr{W}}}\frac{\partial^2 f}{\partial x_i\,\partial x_j}(\theta(t_1),\dots,\theta(t_n))$$
$$+\sum_{i=1}^{n}\mathscr{L}^{\mathscr{W}}(\theta(t_i))\frac{\partial f}{\partial x_i}(\theta(t_1),\dots,\theta(t_n)).$$

Since $\theta(t_i) \in Z^{(1)}$, $\mathscr{L}^{\mathscr{W}}\theta(t_i) = -1/2\theta(t_i)$. At the same time, if $t_i < t_j$, then $\langle \theta(t_i), \theta(t_j)\rangle_{\mathscr{L}\mathscr{W}} = \langle \theta(t_i), \theta(t_i)\rangle_{\mathscr{L}\mathscr{W}} + \langle \theta(t_i), \theta(t_j) - \theta)t_i)\rangle_{\mathscr{L}\mathscr{W}}$ and, since $\mathscr{L}^{\mathscr{W}}((\theta(t_i)(\theta(t_j) - \theta(t_i))) = \theta(t_i)\,\mathscr{L}^{\mathscr{W}}(\theta(t_j) - \theta(t_i)) + (\theta(t_j) - \theta(t_i))\,\mathscr{L}^{\mathscr{W}}(\theta(t_i))$, $\langle \theta(t_i), \theta(t_j) - \theta(t_i)\rangle_{\mathscr{L}\mathscr{W}} = 0$. Finally, $\theta(t)^2 = 2\int_0^t \theta(s)(d\theta(s) + t$; and, because $\int_0^t \theta(s)\,d\theta(s) \in Z^{(2)}$ while $t \in Z^{(0)}$, we find that $\mathscr{L}^{\mathscr{W}}(\theta(t)^2) = -2\int_0^t \theta(s)\,d\theta(s) = -\theta(t)^2 + t$. Hence since $\theta(t)\,\mathscr{L}^{\mathscr{W}}(\theta(t)) = -1/2\theta(t)^2$:

$$\langle \theta(t), \theta(t)\rangle_{\mathscr{L}\mathscr{W}} = t. \tag{3.12}$$

Combining these observations, we now see that

$$\mathscr{L}^{\mathscr{W}}(f(\theta(t_1),..., \theta(t_n))) = 1/2 \sum_{i,j=1}^{n} t_i \wedge t_j \frac{\partial^2 f}{\partial x_i\,\partial x_j}(\theta(t_1),..., \theta(t_n))$$
$$-\frac{1}{2}\sum_{i=1}^{n} \theta(t_i)\frac{\partial f}{\partial x_i}(\theta(t_1),..., \theta(t_n)),$$

which is of course the expression which we would expect on the basis of (3.11).

(3.13) *Remark.* There is a natural "Poincaré inequality" inherent in this calculus. Namely, given $\Phi \in \mathrm{Dom}(\mathscr{L}^{\mathscr{W}})$, observe that

$$-E^{\mathscr{W}}[\Phi\mathscr{L}^{\mathscr{W}}\Phi] = 1/2\sum_{0}^{\infty} nE^{\mathscr{W}}[(E_{Z_{(n)}}\Phi)^2] \geqslant 1/2E^{\mathscr{W}}[(\Phi - E^{\mathscr{W}}[\Phi])^2].$$

Thus

$$E^{\mathscr{W}}[(\Phi - E^{\mathscr{W}}[\Phi])^2] \leqslant E^{\mathscr{W}}[\langle\Phi\rangle^2_{\mathscr{L}\mathscr{W}}], \qquad \Phi \in D(\mathscr{L}^{\mathscr{W}}). \tag{3.14}$$

There is no problem extending these considerations to multi-dimensional Wiener spaces. That is, for $d > 1$, let $\Theta^d = \{\theta \in C([0, \infty), R^d)\colon \theta(0) = 0\}$ and define $\mathscr{B}^d$, $\mathscr{B}^d_t$, and $(\mathscr{B}^d)^t$ accordingly, as in the case $d = 1$. Since Θ^d can be thought of as the d-fold product of Θ with itself, Theorem (2.1) tells us how to introduce a symmetric diffusion semigroup on $L^2(\Theta^d, \mathscr{B}^d, \mathscr{W}^d)$ by taking tensor products. The properties of the multidimensional operations can be read off from those of their one dimensional analogues with the aid of Theorem (2.1).

(3.15) *Warning.* From now on we will drop the superscripts which we have been using to identify the semigroup being discussed. The reason for our deleting them is that the only semigroups with which we will be dealing is the Ornstein–Uhlenbeck semigroup on Wiener space and its higher dimensional analogues. For simplicity, we will refer to the calculus determined by these operations as the *Malliavin Calculus*. Also, unless there is danger of ambiguity, we will not make explicit mention each time of the dimension of the Wiener space in which we are working.

4. The Malliavin Calculus and Stochastic Integrals

We continue with the notation introduced at the end of the preceding section.

Given a $\mathscr{B}_t$-measurable $\alpha \in (\mathrm{Dom}(\mathscr{L}))^d$ and $h > 0$, we have, from Theorem (3.9), that $\alpha \cdot \Delta_k \theta(t) \in \mathrm{Dom}(\mathscr{L})$ and that $\mathscr{L}(\alpha \cdot \Delta_h \theta(t)) = (\mathscr{L}\alpha - 1/2\alpha) \cdot \Delta_h \theta(t)$, where $\Delta_h \theta(t) \equiv \theta(t+h) - \theta(t)$. Hence if α: $[0, \infty) \times \Theta \to R^d$ is a progressively measurable function which is simple (i.e., $\alpha(t) = \alpha([Nt]/N)$, $t \geqslant 0$, for some $N \geqslant 1$) then for all $T > 0$: $\int_0^T \alpha(t) \cdot d\theta(t) \in \mathrm{Dom}(\mathscr{L})$ and

$$\mathscr{L}\left(\int_0^T \alpha(t) \cdot d\theta(t)\right) = \int_0^T (\mathscr{L}\alpha(t) - 1/2\alpha(t)) \cdot d\theta(t). \tag{4.1}$$

In order to get beyond simple integrands, we need the following rather technical approximation result.

(4.2) Lemma. *Let* α: $[0, T] \times \Theta \to R^1$ *be a progressively measurable function such that* $\alpha(t) \in \mathscr{K}_{(2)}(\mathscr{L})$ $(\alpha(t) \in \mathscr{K}_{(4)}(\mathscr{L}))$ *for all* $t \in [0, T]$. *Then* $t \to \|\alpha(t)\|_{\mathscr{K}_{(2)}(\mathscr{L})}$ $(t \to \|\alpha(t)\|_{\mathscr{K}_{(4)}(t)})$ *is a measurable function of* $t \in [0, T]$. *Next, assume that* $\int_0^T \|\alpha(t)\|^2_{\mathscr{K}_{(2)}(\mathscr{L})}\, dt < \infty$ $(\int_0^T \|\alpha(t)\|^4_{\mathscr{K}_{(4)}(\mathscr{L})}\, dt < \infty)$. *Then there is a progressively measurable* β: $[0, T] \times \Theta \to R^1$ *(and a progressively measurable* γ:$[0, T] \times \Theta \to R^1$*) such that for a.e.* $t \in [0, T]$ $\beta(t) = \mathscr{L}\alpha(t)$ *(a.s.,* $\mathscr{W}$*) (and for a.e.* $t \in [0, T]$ $\gamma(t) = \langle \alpha(t), \alpha(t)\rangle$*). Moreover, there exist simple progressively measurable functions* α_n: $[0, T] \times \Theta \to R^1$ *such that* $\alpha_n(t) \in \mathscr{K}_{(2)}(\mathscr{L})$, $t \in [0, T]$ $(\alpha_n(t) \in \mathscr{K}_{(4)}(\mathscr{L})$, $t \in [0, T])$ *and* $\int_0^T (\|\alpha_n(t) - \alpha(t)\|^2_{L^2(\mathscr{W})} + \|\mathscr{L}\alpha_n(t) - \beta(t)\|^2_{L^2(\mathscr{W})})\, dt \to 0$ $(\int_0^T (\|\alpha_n(t) - \alpha(t)\|^4_{L^4(\mathscr{W})} + \|\mathscr{L}\alpha_n(t) - \beta(t)\|^4_{L^4(\mathscr{W})} + \|\langle \alpha_n(t)\rangle - \gamma(t)^{1/2}\|^4_{L^4(\mathscr{W})})\, dt \to 0)$.

Proof. It is clear that $\|\alpha(t)\|_{\mathscr{K}_{(p)}(\mathscr{L})}$ is measurable with respect to $t \in [0, T]$. We now show that if $\int_0^T \|\alpha(t)\|_{\mathscr{K}_{(2)}(\mathscr{L})}\, dt < \infty$ then there is a progressively measurable β: $[0, T] \times \Theta \to R^1$ such that $\beta(t) = \mathscr{L}\alpha(t)$ (a.s., $\mathscr{W}$) for a.e. $t \in [0, T]$. To this end, set $\alpha(t) = 0$ for $t > T$ and choose simple progressively measurable functions $\alpha_n(\cdot)$ so that $\int_0^\infty \|\alpha_n(t) - \alpha(t)\|^2_{L^2(\mathscr{W})}\, dt \to 0$. For $\tau > 0$, set $\beta_n(\tau, t) = T_\tau \alpha(t)$. Then: $\int_0^\infty \|\beta_n(\tau, t) - T_\tau \alpha(t)\|^2_{L^2(\mathscr{W})}\, dt \to 0$ as $n \to \infty$ for each $\tau > 0$. Hence we can find for each $\tau > 0$ a subsequence $\{n_m\}$ such that $\beta_{n_m}(\tau, \cdot) \to T_\tau \alpha(\cdot)$ (a.e., Leb $\times \mathscr{W}$). We now define $\beta(\tau, \cdot) = \overline{\lim}_{m \to \infty} \beta_{n_m}(\tau, \cdot)$ for each $\tau > 0$. It is then clear that if $\beta_N \equiv \beta(1/N)$, $N \geqslant 0$, there is a Lebesgue null set Λ in $[0, \infty)$ such that $\beta_N(t) = T_{1/N}\alpha(t)$ (a.s., $\mathscr{W}$) for $t \notin \Lambda$ and $N \geqslant 1$. In particular,

$$\lim_{N \nearrow \infty} \|N(\beta_N(t) - \alpha(t)) - \mathscr{L}\alpha(t)\|_{L^2(\mathscr{W})} = 0$$

and

$$\sup_{N\geqslant 1} \|N(\beta_N(t)-\alpha(t))\|_{L^2(\mathscr{W})} \leqslant \|\mathscr{L}\alpha(t)\|_{L^2(\mathscr{W})}$$

for all $t \notin \Lambda$. Hence $\int_0^\infty \|N(\beta_N(t)-\alpha(t)) - \mathscr{L}\alpha(t)\|^2_{L^2(\mathscr{W})}\,dt \to 0$ as $N \nearrow \infty$, and therefore we can find a subsequence N_m such that $N_m(\beta_{N_m}(\cdot)-\alpha(\cdot)) - \mathscr{L}\alpha(\cdot) \to 0$ (a.e., Leb $\times \mathscr{W}$). Clearly $\beta(\cdot) = \overline{\lim}_{m\to\infty} N_m(\beta_{N_m}(\cdot)-\alpha(\cdot))$ has the desired properties.

Continuing with the same hypotheses, we next want to construct simple approximants $\alpha_n(\cdot)$ so that $\alpha_n(t) \in \mathrm{Dom}(\mathscr{L})$, $n \geqslant 1$ and $t \geqslant 0$, so that $\int_0^\infty (\|\alpha(t)-\alpha_n(t)\|^2_{L_2(\mathscr{W})} + \|\beta(t) - \mathscr{L}\alpha_n(t)\|_{L^2(\mathscr{W})})\,dt \to 0$ as $n \to \infty$. To this end, define

$$\alpha^{(N)}(t) = \int_0^\infty \rho_N(t-s)\,\alpha(s)\,ds$$

and

$$\beta^{(N)}(t) = \int_0^\infty \rho_N(t-s)\,\beta(s)\,ds,$$

where $\rho \in C_0^\infty((0,\infty))$ satisfies $\int \rho(t)\,dt = 1$ and $\rho_N(t) = N\rho(Nt)$. Then it is easy to check that

$$\lim_{n\to\infty} \int_0^\infty (\|\alpha^{(N)}([nt]/n) - \alpha^{(N)}(t)\|^2_{L^2(\mathscr{W})} + \|\beta^{(N)}([nt]/n) - \beta^{(N)}(t)\|^2_{L^2(\mathscr{W})})\,dt = 0$$

for each $N \geqslant 1$ and that

$$\lim_{N\to\infty} \int_0^\infty (\|\alpha^{(N)}(t) - \alpha(t)\|^2_{L^2(\mathscr{W})} + \|\beta^{(N)}(t) - \beta(t)\|^2_{L^2(\mathscr{W})})\,dt = 0.$$

Thus we will be done once we check that $\alpha^{(N)}(t) \in \mathrm{Dom}(\mathscr{L})$, $t \geqslant 0$, and that $\mathscr{L}\alpha^{(N)}(t) = \beta^{(N)}(t)$, $t \geqslant 0$, for each $N \geqslant 1$. But if $\Phi \in L^2(\mathscr{W})$, then:

$$\begin{aligned}
&E^{\mathscr{W}}[1/\tau(T_\tau(\rho_N * \alpha(t)) - \rho_N * \alpha(t))\Phi] \\
&\quad = (1/\tau)\,E^{\mathscr{W}}[\rho_N * \alpha(t)(T_\tau\Phi - \Phi)] \\
&\quad = \int_0^\infty \rho_N(t-s)\,E^{\mathscr{W}}[(T_\tau\alpha(s) - \alpha(s))/\tau\Phi]\,ds \\
&\to \int_0^\infty \rho_N(t-s)\,E^{\mathscr{W}}[\beta(s)\,\Phi]\,ds = E^{\mathscr{W}}[\beta^{(N)}(t)\,\Phi].
\end{aligned}$$

Since the weak and strong domains of $\mathscr{L}$ coincide, this proves the required property.

Finally, to complete the proof, note that if $\int_0^T \|\alpha(t)\|^4_{\mathscr{T}_{(4)}(\mathscr{L})}\, dt < \infty$, then, by the preceding, $\mathscr{L}\alpha(\cdot)$ and $\mathscr{L}\alpha^2(\cdot)$ admit progressively measurable versions. Hence $\beta(\cdot)$ and $\gamma(\cdot)$ exist. Once this has been established, the existence of simple approximants with the desired properties is proved with precisely the same technique as we just used above. Q.E.D.

(4.3) *Warning.* Given a progressively measurable α: $[0, \infty) \times \Theta \to R^1$ satisfying $\int_0^T \|\alpha(t)\|^4_{\mathscr{T}_{(4)}(\mathscr{L})}\, dt < \infty$, we will use $\mathscr{L}\alpha(t)$ and $\langle \alpha(t), \alpha(t)\rangle$, $t \in [0, T]$, to denote the progressively measurable versions $\beta(t)$ and $\gamma(t)$ just discussed. Clearly this mild abuse of notation causes no problems if $t \to \mathscr{L}\alpha(t)$ and $t \to \langle \alpha(t), \alpha(t)\rangle$ are continuous or if $\mathscr{L}\alpha(\cdot)$ and $\langle \alpha(\cdot), \alpha(\cdot)\rangle$ are integrands in dt or $d\theta(t)$ integrals.

Before proceeding, we introduce a little more notation. Given $\Phi \in (\mathscr{K}_{(p)}(\mathscr{L}))^N$, define

$$|||\Phi|||_{(p)} = E^{\mathscr{W}}\left[\left(\sum_{n=1}^{N} (\Phi_n^2 + (\mathscr{L}\Phi_n)^2 + \langle \Phi_n\rangle^2)\right)^{p/2}\right]^{1/p}. \tag{4.4}$$

Note that since $|\langle \Phi_m, \Phi_n\rangle| \leqslant \langle \Phi_m\rangle\langle \Phi_n\rangle$, $\|\langle \Phi_m, \Phi_n\rangle\|_{L^{p/2}(\mathscr{W})} \leqslant |||\Phi|||^2_{(p)}$. Next, given $T > 0$ and a measurable Φ: $[0, T] \times \Theta \to R^N$ such that $t \to \Phi(t)$ is (a.s., $\mathscr{W}$)-continuous and $\Phi(t) \in (\mathscr{K}_{(p)}(\mathscr{L}))^N$, $t \in [0, T]$, define

$$|||\Phi(\cdot)|||_{(p),T} = E^{\mathscr{W}}\left[\sup_{0 \leqslant t \leqslant T}\left(\sum_{n=1}^{N} (\Phi_n^2(t) + (\mathscr{L}\Phi_n(t))^2 + \langle \Phi_n(t)\rangle^2)\right)^{p/2}\right]^{1/p}. \tag{4.5}$$

Finally, if $\Phi \in (\mathrm{Dom}(\mathscr{L}))^N$, define $\langle\langle \Phi, \Phi\rangle\rangle = ((\langle \Phi_k, \Phi_l\rangle))_{1 \leqslant k,l \leqslant N}$.

(4.6) THEOREM. *Let α: $[0, \infty) \times \Theta \to R^d$ and β: $[0, \infty) \times \Theta \to R^1$ be progressively measurable functions such that $\alpha(t) \in (\mathscr{K}_{(4)}(\mathscr{L}))^d$ and $\beta(t) \in \mathscr{K}_4(\mathscr{L})$ for all $t \geqslant 0$. Assume that $\int_0^T (|||\alpha(t)|||^4_{(4)} + |||\beta(t)|||^4_{(4)})\, dt < \infty$ for all $T > 0$ and set $\xi(T) = \int_0^T \alpha(t) \cdot d\theta(t) + \int_0^T \beta(t)\, dt$, $T \geqslant 0$. Then $\xi(T) \in \mathscr{K}_{(4)}(\mathscr{L})$ for all $T \geqslant 0$, $\mathscr{L}(\xi(T))$ is given by*

$$\mathscr{L}(\xi(T)) = \int_0^T (\mathscr{L}(\alpha(t)) - 1/2\alpha(t)) \cdot d\theta(t) + \int_0^T \mathscr{L}(\beta(t))\, dt, \qquad T \geqslant 0, \tag{4.7}$$

there is for each $p \in [4, \infty)$ a non-decreasing C_p: $[0, \infty) \to [0, \infty)$ such that

$$|||\xi(\cdot)|||^p_{(p),T} \leqslant C_p(T) \int_0^T (|||\alpha(t)|||^p_{(p)} + |||\beta(t)|||^p_{(p)})\, dt, \tag{4.8}$$

and $\langle \xi(T), \xi(T)\rangle$ *is given by*

$$\langle \xi(T), \xi(T)\rangle = 2\int_0^T \langle \xi(t), \alpha(t)\rangle \cdot d\theta(t)$$

$$+\int_0^T (2\langle \xi(t), \beta(t)\rangle + |\alpha(t)|^2 + \langle \alpha(t)\rangle^2)\, dt, \qquad T > 0, \tag{4.9}$$

where $\langle \xi(t), \alpha(t)\rangle \equiv (\langle \xi(t), \alpha_1(t)\rangle, ..., \langle \xi(t), \alpha_d(t)\rangle)$ *and* $\langle \alpha(t)\rangle \equiv (\sum_1^d \langle \alpha_k(t)\rangle^2)^{1/2}$.

Proof. First assume that $\alpha(\cdot)$ and $\beta(\cdot)$ are simple. Then, by (4.1), it is easy to see that $\xi(T) \in \mathrm{Dom}(\mathscr{L})$, $T > 0$, and that $\mathscr{L}(\xi(T))$ is given by (4.7). Moreover, using Burkholder's inequality, we derive from (4.7)

$$E^{\mathscr{W}}\left[\sup_{0 \leqslant t \leqslant T} |\mathscr{L}(\xi(t))|^p\right] \leqslant C_p \left(T^{p/2-1}\int_0^T (\|\mathscr{L}(\alpha(t))\|_{L^p(\mathscr{W})}^p + \|\alpha(t)\|_{L^p(\mathscr{W})}^p)\, dt\right.$$

$$\left. + T^{p-1}\int_0^T \|\beta(t)\|_{L^p(\mathscr{W})}^p\, dt\right), \tag{4.10}$$

for $p \in [2, \infty)$. Using (4.10) with $p = 2$ and applying Lemma (4.2), we see that if $\alpha: [0, \infty) \times \Theta \to R^d$ and $\beta: [0, \infty) \times \Theta \to R^1$ are progressively measurable functions satisfying $\alpha(t) \in (\mathrm{Dom}(\mathscr{L}))^d$ and $\beta(t) \in \mathrm{Dom}(\mathscr{L})$, $t \geqslant 0$, with $\int_0^T (|||\alpha(t)|||_{(2)}^2 + |||\beta(t)|||_{(2)}^2)\, dt < \infty$, $T > 0$, then $\xi(T) \in \mathrm{Dom}(\mathscr{L})$, $T \geqslant 0$, $\mathscr{L}(\xi(T))$ is given by (4.7), and (4.10) holds.

We next assume that $\alpha(\cdot)$ and $\beta(\cdot)$ are simple functions satisfying our hypotheses. From Theorem (3.9), it is easy to see that $\sup_{0 \leqslant t \leqslant T} |||\xi(t)|||_{(4)} < \infty$ for all $T > 0$. Hence:

$$\int_0^T (|||\xi(t)\, \alpha(t)|||_{(2)}^2 + |||\xi(t)\, \beta(t)|||_{(2)}^2 + |||\, |\alpha(t)|^2\, |||_{(2)}^2)\, dt < \infty$$

for all $T > 0$. Since

$$\xi^2(T) = 2\int_0^T \xi(t)\, \alpha(t) \cdot d\theta(t) + \int_0^t (2\xi(t)\, \beta(t) + |\alpha(t)|^2)\, dt,$$

$$T > 0,$$

we can now apply the result proved in the preceding paragraph to conclude that $\xi^2(T) \in \mathrm{Dom}(\mathscr{L})$, $T > 0$, and that

$$\mathscr{L}(\xi^2(T)) = \int_0^T (2\mathscr{L}(\xi(t)\, \alpha(t)) - \xi(t)\, \alpha(t)) \cdot d\theta(t)$$

$$+\int_0^T (2\mathscr{L}(\xi(t)\, \beta(t)) + \mathscr{L}(|\alpha(t)|^2))\, dt.$$

At the same time

$$\xi(T)\,\mathscr{L}(\xi(T))=\int_0^T (\xi(t)(\mathscr{L}(\alpha(t))-1/2\alpha(t))$$
$$+(\mathscr{L}\xi(t)\,\alpha(t)))\cdot d\theta(t)$$
$$+\int_0^T (\xi(t)\,\mathscr{L}(\beta(t))+(\mathscr{L}(\xi(t))\,\beta(t)$$
$$+\alpha(t)\cdot(\mathscr{L}(\alpha(t))-1/2\alpha(t)))\,dt.$$

Combining these, we see that $\langle\xi(T),\xi(T)\rangle$ is given by (4.9). Moreover, (4.8) is a consequence of (4.10), (4.9), Burkholder's inequality, and $|\langle\xi(t),\alpha(t)\rangle|\leqslant\langle\xi(t)\rangle^2+\langle\alpha(t)\rangle^2$.

Finally, the restriction to simple $\alpha(\cdot)$ and $\beta(\cdot)$ can now be removed by an application of Lemma (4.2) in conjunction with (4.8). Q.E.D.

We now want to apply these results to solutions of stochastic integral equations. Our next theorem deals with a somewhat more general situation than we need immediately but its full generality will be useful later on.

(4.11) THEOREM. *Let* σ: $[0,\infty)\times R^M\times R^N\to R^N\otimes R^d$ *and* b: $[0,\infty)\times R^M\times R^N\to R^N$ *be measurable functions which are twice continuously differentiable with respect to* $(\eta,y)\in R^M\times R^N$ *for each* $t\in[0,\infty)$. *Assume in addition that for* $\alpha\in(\{0,1,2\})^M$ *and* $\beta\in(\{0,1,2\})^N$ *satisfying* $|\alpha|+|\beta|\leqslant 2$ *there exist* $C(\alpha,\beta)<\infty$ *and* $\gamma(\alpha,\beta)$, $\lambda(\alpha,\beta)\in[0,\infty)$ *such that* $\gamma(\alpha,\beta)=\lambda(\alpha,\beta)=0$ *when* $|\beta|=1$ *and* $\|D^\alpha_\eta D^\beta_y\sigma(t,\eta,y)\|_{H.S.}\vee|D^\alpha_\eta D^\beta_y\, b(t,\eta,y)|\leqslant C(\alpha,\beta)(1+|\eta|^{\gamma(\alpha,\beta)}+|y|^{\lambda(\alpha,\beta)})$ *for* $(t,\eta,y)\in[0,\infty)\times R^M\times R^N$. *Finally, let* η: $[0,\infty)\times\Theta\to R^N$ *be a progressively measurable function satisfying* $\eta(t)\in(\mathscr{K}(\mathscr{L}))^M$, $t\geqslant 0$, *and* $\int_0^T|||\eta(t)|||^p_{(p)}\,dt<\infty$, $T>0$ *and* $p\in[4,\infty)$. *Then for each* $y\in R^N$, *there is an (a.s.,* $\mathscr{W}$*) unique progressively measurable* y: $[0,\infty)\times\Theta\to R^N$ *satisfying*:

$$y(T)=y+\int_0^T\sigma(t,\eta(t),y(t))\,d\theta(t)+\int_0^T b(t,\eta(t),y(t))\,dt,\qquad T>0. \tag{4.12}$$

Moreover: $y(T)\in(\mathscr{K})^N$ *for all* $T\geqslant 0$; $T\to(y(T),\langle\langle y(T),y(T)\rangle\rangle,\mathscr{L}(y(T)))$ *is (a.s.,* $\mathscr{W}$*) continuous; and* $|||\,y(\cdot)|||_{(p),T}<\infty$, $T\geqslant 0$ *and* $p\in[4,\infty)$.

Proof. It is an immediate consequence of Itô's method that (4.12) admits only one solution $y(\cdot)$ and that $E^{\mathscr{W}}[\sup_{0\leqslant t\leqslant T}|y(t)-y^{(\nu)}(t)|^p]\to 0$ as $\nu\to\infty$ for all $T>0$ and $p\in[1,\infty)$, where $y^{(0)}(\cdot)\equiv y$ and

$$y^{(\nu)}(T)=y+\int_0^T\sigma(t,\eta(t),y^{(\nu-1)}(t))\,d\theta(t)$$

$$+\int_0^T b(t,\eta(t),y^{(\nu-1)}(t))\,dt,\qquad T\geqslant 0,$$

for $\nu\geqslant 1$. By Theorem (4.6) and induction, it is clear that $y^{(\nu)}(T)\in(\mathscr{K}(\mathscr{L}))^N$ and $|||\,y^{(\nu)}(\cdot)|||_{(p),T}<\infty$ for all $\nu\geqslant 0$, $T\geqslant 0$, and $p\in[4,\infty)$. Thus, if we prove that $\sup_\nu|||\,y^{(\nu)}(\cdot)|||_{(p),T}<\infty$ for all $T\geqslant 0$ and $p\in[4,\infty)$, then we will know that $y(T)\in(\mathscr{K}(\mathscr{L}))^N$ and $\sup_{0\leqslant t\leqslant T}|||\,y(t)|||_{(p)}<\infty$ for all $T\geqslant 0$ and $p\in[4,\infty)$. In particular, after another application of Theorem (4.6), it will follow from (4.12) that $T\to(y(T),\langle\langle y(T),y(T)\rangle\rangle,\ \mathscr{L}(y(T)))$ is continuous (a.s., $\mathscr{W}$) and that $|||\,y(\cdot)|||_{(p),T}<\infty$ for all $T\geqslant 0$ and $p\in[4,\infty)$. It therefore suffices for us to prove that $\sup_\nu|||\,y^{(\nu)}(\cdot)|||_{(p),T}<\infty$ for all $T\geqslant 0$ and $p\in[4,\infty)$.

We already know that $\sup_\nu\|\sup_{0\leqslant t\leqslant T}y^{(\nu)}(t)\|_{L^2(\mathscr{W})}<\infty$ for all $T\geqslant 0$ and $p\in[1,\infty)$. To prove that $\sup_\nu\|\sup_{0\leqslant t\leqslant T}\langle y^{(\nu)}(t)\rangle\|_{L^2(\mathscr{W})}<\infty$ for all $T\geqslant 0$ and $p\in[4,\infty)$, we use (4.9) to derive the existence for each $p\in[2,\infty)$ of a non-decreasing $C_p\colon[0,\infty)\to[0,\infty)$ such that:

$$\|\sup_{0\leqslant t\leqslant T}\langle y^{(\nu+1)}(t)\rangle^2\|_{L^p(\mathscr{W})}^p\leqslant C_p(T)\int_0^T\Bigg(\sum_{j=1}^N\sum_{k=1}^d\|\langle\sigma_k^j(t,\eta(t),y^\nu(t))\rangle^2\|_{L^p(\mathscr{W})}^p$$

$$+\sum_{j=1}^N\|\langle b^j(t,\eta(t),y^\nu(t))\rangle^2\|_{L^p(\mathscr{W})}^p$$

$$+\|\operatorname{Trace}\sigma\sigma^*(t,\eta(t),y^{(\nu)}(t))\|_{L^p(\mathscr{W})}^p\Bigg)\,dt.$$

Since $\sup_\nu\|\sup_{0\leqslant t\leqslant T}|y^{(\nu)}(t)|\|_{L^p(\mathscr{W})}<\infty$ for all $T\geqslant 0$, it is clear that the third term in the integrand can be bounded independent of ν. To handle the other two terms, let $F\colon[0,\infty)\times R^M\times R^N\to R^1$ be a function satisfying the hypotheses placed on $\sigma(\cdot)$ and $b(\cdot)$. Then $\langle F(t,\eta(t),y^\nu(t))\rangle^2$ is a finite linear combination of terms of the following three types:

$$\text{(I)}\quad X^{(\nu)}(t)\equiv\frac{\partial F}{\partial\eta_\lambda}\frac{\partial F}{\partial\eta_{\lambda'}}(t,\eta(t),y^{(\nu)}(t))\langle\eta_\lambda(t),\eta_{\lambda'}(t)\rangle,$$

$$\text{(II)}\quad Y^{(\nu)}(t)\equiv\frac{\partial F}{\partial\eta_\lambda}\frac{\partial F}{\partial y_l}(t,\eta(t),y^{(\nu)}(t))\langle\eta_\lambda(t),y_l^{(\nu)}(t)\rangle,$$

and

$$\text{(III)}\quad Z^{(\nu)}(t) \equiv \frac{\partial F}{\partial y_l}\frac{\partial F}{\partial y_{l'}}(t, \eta(t), y^{(\nu)}(t))\langle y_l^{(\nu)}(t), y_{l'}^{(\nu)}(t)\rangle.$$

It is easy to see from our preceding considerations that $\sup_\nu \int_0^T \|X^{(\nu)}(t)\|^p_{L^p(\mathscr{W})}\,dt < \infty$, $T \geqslant 0$ and $p \in [2, \infty)$. Also, since $|\langle \eta_\lambda(t), y_l(t)\rangle| \leqslant \langle \eta_\lambda(t)\rangle\langle y_l(t)\rangle$: $|Y^{(\nu)}(t)| \leqslant \tilde{Y}^{(\nu)}(t) + \langle y_l^\nu(t)\rangle^2$, where $\sup_\nu \int_0^T \|\tilde{Y}^{(\nu)}(t)\|^p_{L^p(\mathscr{W})}\,dt < \infty$, $T \geqslant 0$ and $p \in [2, \infty)$. Finally, $|Z^\nu(t)| \leqslant C\langle y_l^{(\nu)}(t)\rangle\langle y_{l'}^{(\nu)}(t)\rangle$. Hence, for each $p \in [2, \infty)$ there exist nondecreasing $A_p\colon [0, \infty) \to [0, \infty)$ and $B_p\colon [0, \infty)$ such that:

$$\|\sup_{0\leqslant t\leqslant T} \langle y^{(\nu+1)}(t)\rangle^2\|_{L^2(\mathscr{W})} \leqslant A_p(T) + B_p(T)\int_0^T \|\langle y^{(\nu)}(t)\rangle^2\|^p_{L^p(\mathscr{W})}\,dt.$$

Using Gromwall's inequality, we conclude from this that

$$\sup_\nu \|\sup_{0\leqslant t\leqslant T} \langle y^{(\nu)}(t)\rangle^2\|^p_{L^p(\mathscr{W})} \leqslant A_p(T)\,e^{B_p(T)T}, \qquad T \geqslant 0 \quad\text{and}\quad p \in [2, \infty).$$

To complete the theorem we must still show that

$$\sup_\nu \|\sup\nolimits_{0\leqslant t\leqslant T} |\mathscr{L}(y^{(\nu)}(t))|\|_{L^p(\mathscr{W})} < \infty \text{ for all } T \geqslant 0 \text{ and } p \in [4, \infty).$$

But, starting from (4.7), we see that for each $p \in [4, \infty)$ there is a nondecreasing $C_p\colon [0, \infty) \to [0, \infty)$ such that

$$\begin{aligned}\|\sup_{0\leqslant t\leqslant T} |\mathscr{L}(y^{(\nu+1)}(t))|\|^p_{L^p(\mathscr{W})} \leqslant C_p(T)\Big(1 + \int_0^T (\|y^{(\nu)}(t)\|^p_{L^p(\mathscr{W})}\\ + \|\langle y^{(\nu)}(t)\rangle^2\|^p_{L^p(\mathscr{W})} + \|\mathscr{L}(y^{(\nu)}(t))\|^p_{L^p(\mathscr{W})})\,dt\Big).\end{aligned}$$

In view of the preceding considerations, it follows from this that

$$\|\sup_{0\leqslant t\leqslant T} |\mathscr{L}(y^{(\nu+1)})(t))|\|^p_{L^p(\mathscr{W})} \leqslant A_p(T) + B_p(T)\int_0^T \|\mathscr{L}(y^{(\nu)}(t))\|^p_{L^p(\mathscr{W})}\,dt,$$

where $A_p\colon [0, \infty) \to [0, \infty)$ and $B_p\colon [0, \infty) \to [0, \infty)$ are non-decreasing. Hence the desired estimate follows by another application of Gromwall's inequality. Q.E.D.

(4.13) COROLLARY. *Let $\sigma\colon [0, \infty) \times R^N \to R^N \otimes R^d$ and $b\colon [0, \infty) \times R^N \to R^N$ be measurable functions which are three times continuously differentiable with respect to $x \in R^N$ for each $t \in [0, \infty)$. Assume that*

$\sup_{t \geqslant 0} \sup_{x \in R^N} |D_x^\alpha \sigma_k^j(t, x)| \vee |D_x^\alpha b^j(t, x)| < \infty$ *for* $1 \leqslant |\alpha| \leqslant 3$. *Then for each* $x \in R^N$ *there is a unique progressively measurable solution* $x(\cdot) = x(\cdot, x)$ *to*

$$x(T) = x + \int_0^T \sigma(t, x(t))\, d\theta(t) + \int_0^T b(t, x(t))\, dt, \qquad T \geqslant 0. \tag{4.14}$$

Furthermore, $x(T) \in (\mathscr{K}(\mathscr{L}))^N$ *for all* $T \geqslant 0$, $T \to (x(T), A(T), \mathscr{L}(x(T)))$ *is continuous (a.s.,* $\mathscr{W}$*), where* $A(T) \equiv \langle\langle x(T), x(T)\rangle\rangle$, *and* $|||x(\cdot)|||_{(p),T} < \infty$ *for all* $T \geqslant 0$ *and* $p \in [4, \infty)$. *Finally, if*

$$S_k(t, x) = \left(\left(\frac{\partial \sigma_k^i}{\partial x_j}(t, x)\right)\right)_{1 \leqslant i,j \leqslant N}, \qquad 1 \leqslant k \leqslant d,$$

$$B(t, x) = \left(\left(\frac{\partial b^i}{\partial x_j}(t, x)\right)\right)_{1 \leqslant i,j \leqslant N},$$

and

$$a(t, x) = \sigma\sigma^*(t, x),$$

then $A(\cdot)$ *is the unique progressively measurable solution to*:

$$\begin{aligned} A(T) = {} & \sum_{k=1}^d \int_0^T \{S_k(t, x(t)), A(t)\}\, d\theta_k(t) \\ & + \int_0^T \Big(\{B(t, x(t)), A(t)\} + \sum_{k=1}^d S_k(t, x(t))\, A(t)\, S_k(t, x(t))^* \\ & + a(t, x(t))\Big)\, dt, \qquad T \geqslant 0, \end{aligned} \tag{4.15}$$

where $\{M_1, M_2\} \equiv M_1 M_2^* + M_2 M_1^*$ *for* $M_1, M_2 \in R^N \otimes R^N$. *In particular,* $A(T) \in (\mathscr{K}(\mathscr{L}))^{N^2}$, $T \geqslant 0$, $T \to (A(T), \langle\langle A(T), A(T)\rangle\rangle, \mathscr{L}(A(T)))$ *is continuous (a.s.,* $\mathscr{W}$*) and* $|||A(\cdot)|||_{(p),T} < \infty$ *for all* $T \geqslant 0$ *and* $p \in [4, \infty)$.

Proof. In view of Theorem (4.11), we know that $x(\cdot)$ is unique, $x(T) \in (\mathscr{K}(\mathscr{L}))^N$ for $T \geqslant 0$, $T \to (x(T), A(T), \mathscr{L}(x(T)))$ is continuous, and $|||x(\cdot)|||_{(p),T} < \infty$ for all $T \geqslant 0$ and $p \in [4, \infty)$. Furthermore, that $A(\cdot)$ satisfies (4.15) can be easily seen from (4.9). To complete the proof, take $\eta(\cdot) = x(\cdot)$ and $y(\cdot) = A(\cdot)$ in Theorem (4.11). Then, by Theorem (4.11), $y(\cdot)$ is uniquely determined by (4.15), $y(T) \in (\mathscr{K}(\mathscr{L}))^{N^2}$ for $T \geqslant 0$, $T \to (y(T), \langle\langle y(T), y(T)\rangle\rangle, \mathscr{L}(y(T)))$ is (a.s., $\mathscr{W}$) continuous, and $|||y(\cdot)|||_{(p),T} < \infty$ for all $T \geqslant 0$ and $p \in [4, \infty)$. Q.E.D.

5. Preliminary Applications; the Forward Variables

Throughout this section we will be working with the following situation: $(\Theta, \mathscr{B}, \mathscr{W})$ is d-dimensional Wiener space and $\mathscr{L}$ and $\langle \cdot, \cdot \rangle$ are the associated Malliavin operations; σ: $[0, \infty) \times R^N \to R^N \otimes R^d$ and b: $[0, \infty) \times R^N \to R^N$ are bounded measurable functions such that $\sigma(t, \cdot)$ and $b(t, \cdot)$ are three times continuously differentiable for each $t \in [0, \infty)$ and $\sup_{t \geqslant 0} \|\sigma(t, \cdot)\|_{C_b^3(R^N)} \vee \|b(t, \cdot)\|_{C_b^3(R^N)} < \infty$; $x(\cdot) = x(t, x)$ is the unique progressively measurable solution to

$$x(T) = x + \int_0^T \sigma(t, x(t))\, d\theta(t) + \int_0^T b(t, x(t))\, dt, \qquad T \geqslant 0. \tag{5.1}$$

By Corollary (4.13), $x(T) \in (\mathscr{K}(\mathscr{L}))^N$ for $T \geqslant 0$; $T \to (x(T), \langle\langle x(T), x(T) \rangle\rangle, \mathscr{L}(x(T)))$ is continuous (a.s., $\mathscr{W}$); and if $A(T) = A(T, x) \equiv \langle\langle x(T, x), x(T, x) \rangle\rangle$, then $A(T) \in (\mathscr{K}(\mathscr{L}))^{N^2}$ and $T \to (A(T), \langle\langle A(T), A(T) \rangle\rangle, \mathscr{L}(A(T)))$ is continuous (a.s., $\mathscr{W}$). Moreover, since we have assumed that $\sigma(\cdot)$ and $b(\cdot)$ are bounded, it is easy to see from the proof of Theorem (4.11) that for each $p \in [4, \infty)$ there is a non-decreasing C_p: $[0, \infty) \to [0, \infty)$ such that

$$\sup_{x \in R^N} |||x(\cdot, x) - x|||_{(p), T} \vee |||A(\cdot, x)|||_{(p), T} \leqslant C_p(T), \qquad T \geqslant 0. \tag{5.2}$$

In order to complete the program outlined in Remark (1.21), it is still necessary for us to investigate when $A(T)$ is non-degenerate. Our investigation will rely on an analysis of Eq. (4.15).

(5.3) Lemma. *Let* U_k: $[0, \infty) \times \Theta \to R^N \otimes R^N$, $1 \leqslant k \leqslant d$, *and* V: $[0, \infty) \times \Theta \to R^N \otimes R^N$ *be bounded progressively measurable functions. Then there is precisely one progressively measurable* X:$[0, \infty) \times \Theta \to R^N \otimes R^N$ *satisfying*

$$X(T) = I + \sum_{k=1}^{d} \int_0^T U_k(t) X(t)\, d\theta_k(t) + \int_0^T V(t) X(t)\, dt, \qquad T \geqslant 0. \tag{5.4}$$

Furthermore, $X(\cdot)$ *is (a.s.,* $\mathscr{W}$*) non-singular and* $X(\cdot)^{-1}$ *is the unique progressively measurable* Y:$[0, \infty) \times \Theta \to R^N \otimes R^N$ *satisfying*

$$\begin{aligned} Y(T) = I - \sum_{k=1}^{d} \int_0^T Y(t)\, U_k(t)\, d\theta_k(t) \\ + \int_0^T \left(\sum_{k=1}^{d} Y(t)\, U_k(t)^2 - Y(t)\, V(t) \right) dt, \qquad T \geqslant 0. \end{aligned} \tag{5.5}$$

In particular, if $X(s,t) \equiv X(t)X(s)^{-1}$ *for* $0 \leqslant s \leqslant t$, *then* $(s,t) \to X(s,t)$ *is continuous* (*a.s.*, $\mathscr{W}$) *and for each* $s \geqslant 0$, $X(s, \cdot \vee s)$ *is uniquely determined by*

$$X(s,T) = I + \sum_{k=1}^{d} \int_s^T U_k(t)\, X(s,t)\, d\theta_k(t) + \int_s^T V(t)\, X(s,t\,(dt, \qquad T \geqslant s. \tag{5.6}$$

Proof. The existence and uniqueness of solutions to (5.4) and (5.5) are easy consequences of elementary Itô theory. Moreover, if $Z(t) \equiv X(t)\, Y(t)$, then $Z(\cdot)$ satisfies:

$$\begin{aligned} Z(T) = I &+ \sum_{k=1}^{d} \int_0^T (U_k(t)\, Z(t) - Z(t)\, U_k(t))\, d\theta_k(t) \\ &+ \int_0^T \left(\sum_{k=1}^{d} (Z(t)\, U_k(t)^2 - U_k(t)^2\, Z(t)) + V(t)\, Z(t) - Z(t)\, V(t) \right) dt \end{aligned}$$

for $T \geqslant 0$. Since this equation has only one solution and the identity is a solution, we conclude that $Z(T) = I$, $T \geqslant 0$. Finally, it is trivial to verify that $X(s, \cdot \vee s)$ satisfies (5.6) and is determined by this equation. Q.E.D.

Define S_k, B, and a as in Corollary (4.13) and define $X(\cdot) = X(\cdot, x)$ by

$$X(T) = I + \sum_{k=1}^{d} \int_0^T S_k(t, x(t))\, X(t)\, d\theta_k(t) + \int_0^T B(t, x(t))\, X(t) dt, \qquad T \geqslant 0, \tag{5.7}$$

Using (4.15) and the standard variation of parameters technique, we arrive at:

$$A(T) = \int_0^T X(s,T)\, a(s, x(s))\, X(s,T)^*\, ds, \qquad T \geqslant 0, \tag{5.8}$$

where $X(s,t) \equiv X(t)\, X(s)^{-1}$, $0 \leqslant s \leqslant t$.

These considerations already allow us to draw some interesting conclusions. Indeed, suppose that $a(\cdot) \geqslant \varepsilon I$ for some $\varepsilon > 0$. Then from (5.8) we obtain

$$A(T)^{-1} \leqslant \frac{1}{\varepsilon T^2} \int_0^T (X(s,T)\, X(s,T)^*)^{-1}\, ds,$$

since one can show that the inverse of a convex combination of symmetric positive definite matrices is dominated by the convex combination of their inverses. From this it is a rather simple matter to conclude that for each $p \in [1, \infty)$ there is a $C_p < \infty$ such that $\|(\det A(T))^{-1}\|_{L^p(\mathscr{W})} \leqslant C_p T^{-N/2}$. Hence, by Remark (1.21), we see that for $T > 0$ the distribution of $x(T)$ under $\mathscr{W}$ admits a bounded Hölder continuous density. Before carrying out such a line of reasoning in detail, we first improve (5.2) in order to sharpen our final estimates.

(5.9) LEMMA. *For each $p \in [1, \infty)$ there is a C_p, depending only on $\sup_{t \geqslant 0}(\|\sigma(t, \cdot)\|_{C_b^2(R^N)} \vee \|b(t, \cdot)\|_{C_b^2(R^N)}$, such that*

$$\max_{1 \leqslant m,n \leqslant N} \|A_{m,n}(T)\|_{L^p(\mathscr{W})} \leqslant C_p T, \qquad 0 \leqslant T \leqslant 1, \tag{5.10}$$

$$\lim_{\substack{1 \leqslant l \leqslant N \\ 1 \leqslant m,n \leqslant N}} \|\langle x_l(T), A_{m,n}(T)\rangle\|_{L^p(\mathscr{W})} \leqslant C_p T^2, \qquad 0 \leqslant T \leqslant 1, \tag{5.11}$$

and

$$\max_{1 \leqslant l \leqslant N} \|\mathscr{L}(x_l(T))\|_{L^p(\mathscr{W})} \leqslant C_p T^{1/2}, \qquad 0 \leqslant T \leqslant 1. \tag{5.12}$$

Proof. To prove (5.10), note that from (5.8) we have:

$$\begin{aligned}\frac{1}{T}\|\mathrm{Trace}(A(T))\|_{L^p(\mathscr{W})} &\leqslant \frac{1}{T}\int_0^T \|\mathrm{Trace}(X(s, T)\, a(s, x(s))\, X(s, T)^*)\|_{L^p(\mathscr{W})}\, ds \\ &\leqslant \Lambda \sup_{0 \leqslant s \leqslant T} \|\mathrm{Trace}(X(s, T)\, X(s, T)^*)\|_{L^p(\mathscr{W})},\end{aligned}$$

where $\Lambda = \sup_{(t,x)} \|a(t, x)\|_{op}$. At the same time, if $Y(s, t) = X(s, t)\, X(s, t)^*$, $t \geqslant s$, then from (5.6):

$$\begin{aligned}Y(s, T) = I + \sum_{k=1}^{d} \int_s^T &\{S_k(t, x(t)), Y(s, t)\}\, d\theta_k(t) \\ + \int_s^T &(\{B(t, x(t)), Y(s, t)\} \\ &+ \sum_{k=1}^{d} (S_k(t, x(t))\, Y(s, t)\, S_k(t, x(t))^*)\, dt\end{aligned} \tag{5.13}$$

for $T \geqslant s$. Using the fact that if $K \in R^N \otimes R^N$ and $M \in R^N \otimes R^N$ is symmetric and non-negative definite then $\operatorname{Trace}(KM) \leqslant \|K\|_{op} \operatorname{Trace}(M)$, we see from (5.13) that

$$\operatorname{Trace}(Y(s, T)) = N + \sum_1^d \int_s^T \gamma_k(t) \operatorname{Trace}(Y(s, t))\, d\theta_k(t) + \int_0^T \beta(t) \operatorname{Trace}(Y(s, t))\, dt, \qquad T \geqslant s,$$

where the $\gamma_k(\cdot)$ and $\beta(\cdot)$ are bounded progressively measurable functions which can be bounded in terms of $\sup_{t \geqslant 0} \|\sigma(t, \cdot)\|_{C_b^1(R^N)} \vee \|b(t, \cdot)\|_{C_b^1(R^N)}$. Thus:

$$\operatorname{Trace} Y(s, T) = N \exp\left[\sum_{k=1}^d \int_s^T \gamma_k(t)\, d\theta_k(t) + \int_0^T \Big(\beta(t) - 1/2 \sum_{k=1}^d \gamma_k^2(t)\Big)\, dt\right];$$

and so for each $p \in [1, \infty)$ there is a $B_p < \infty$ depending only on $\sup_{t \geqslant 0} \|\sigma(t, \cdot)\|_{C_b^1(R^N)} \vee \|b(t, \cdot)\|_{C_b^1(R^N)}$ such that:

$$\Big\|\sup_{s \leqslant t \leqslant T} \operatorname{Trace}(X(s, T)\, X(s, T)^*)\Big\|_{L^p(\mathscr{W})} \leqslant N e^{B_p(T-s)}, \qquad T \geqslant s. \tag{5.14}$$

Combining this with the preceding, we now see that:

$$\|\operatorname{Trace} A(T)\|_{L^p(\mathscr{W})} \leqslant A N T e^{B_p T}, \qquad p \in [1, \infty) \quad \text{and} \quad T \geqslant 0. \tag{5.15}$$

Since $|A_{m,n}(T)| \leqslant (A_{mm}(T) + A_{nn}(T))/2$, we have now proved (5.10).

Turning to (5.11), observe that:

$$\begin{aligned}\langle x_l(T), A(T)_{m,n}\rangle = {} & \sum_{k=1}^d \int_0^T [\langle x_l(t), \{S_k(t), A(t)\}_{m,n}\rangle + \langle \sigma_l^k(t), A(t)_{m,n}\rangle]\, d\theta_k(t) \\ & + \int_0^T [\langle x_l(t), \{B(t), A(t)\}_{m,n}\rangle \\ & + \langle x_l(t), a_{m,n}(t)\rangle + \langle b_l(t), A(t)_{m,n}\rangle \\ & + \sum_{k=1}^d (\langle x_l(t), (S_k(t)\, A(t)\, S_k(t)^*)_{m,n}\rangle \\ & + \langle \sigma_l^k(t), \{S_k(t), A(t)\}_{m,n}\rangle \\ & + \sigma_l^k(t) \{S_k(t), A(t)\}_{m,n})]\, dt,\end{aligned}$$

where $S_k(t) = S_k(t, x(t))$, $\sigma(t) = \sigma(t, x(t))$, $B(t) = B(t, x(t))$, $a(t) = a(t, x(t))$, and $b(t) = b(t, x(t))$. Thus, if $Y_\mu(t) = \langle x_{\mu_1}(t) = \langle x_{\mu_1}(t), A(t)_{\mu_1, \mu_2}\rangle$ for $\mu \in (\{1,..., N\})^3$, then

$$Y_\mu(T) = \sum_{k=1}^{d} \int_0^T \left(\sum_\nu \gamma_\nu^k(t)\, Y_\nu(t) + \beta^k(t)\right) d\theta_k(t)$$
$$+ \int_0^T \left(\sum_\nu \gamma_\nu^0(t) + \beta^0(t)\right) dt,$$

where the $\gamma_\nu^k(\cdot)$'s an $\beta^k(\cdot)$'s are progressively measurable functions satisfying the following bounds:

$$\max_{\substack{a \leqslant k \leqslant d \\ \mu \in (\{1,...,N\})^3}} |\gamma_\mu^k(\cdot)| \leqslant K,$$

$$\max_{1 \leqslant k \leqslant d} |\beta^k(\cdot)| \leqslant K(\mathrm{Trace}(A(\cdot)))^2,$$

$$|\beta^0(\cdot)| \leqslant K(\mathrm{Trace}(A(\cdot)) + (\mathrm{Trace}\, A(\cdot))^2),$$

where $K < \infty$ depends only on $\sup_{t \geqslant 0} \|\sigma(t, \cdot)\|_{C_b^2(R^N)} \vee \|b(t, \cdot)\|_{C_b^2(R^N)}$. Since from the equation above plus Burkholder's inequality

$$E^{\mathscr{W}}[| Y_\mu(T)|^p] \leqslant C_p E^{\mathscr{W}} \left[\left(\int_0^T \sum_{k=1}^{d} \left(\sum_\nu \gamma_\nu^k(t)\, Y_\nu(t) + \beta^k(t)\right)^2\right)^{p/2}\right]$$
$$+ C_p E^{\mathscr{W}} \left[\left|\int_0^T \left(\sum_\nu \gamma_\nu^0(t)\, Y_\nu(t) + \beta^0(t)\right) dt\right|^p\right]$$

for each $p \in [2, \infty)$, we see that if $\bar{Y}(\cdot) = (\sum_\mu Y_\mu(\cdot)^2)^{1/2}$ then for each $p \in [2, \infty)$ there is a $K_p < \infty$, depending only on the K above, such that:

$$E^{\mathscr{W}}[| \bar{Y}(T)|^p] \leqslant K_p(T^{p/2-1} \int_0^T E^{\mathscr{W}}[| \bar{Y}(t)|^p]\, dt$$
$$+ T^{p/2-1} \int_0^T E^{\mathscr{W}}[(\mathrm{Trace}(A(t)))^{2p}]\, dt$$
$$+ T^{p-1} \int_0^T E^{\mathscr{W}}[(\mathrm{Trace}(A(t)) + (\mathrm{Trace}(A(t)))^2)^p]\, dt),$$
$$T \geqslant 0.$$

Hence, by (5.15), we now see that for $p \in [2, \infty)$ there exist $A_p < \infty$ and $B_p < \infty$, depending only on $\sup_{t \geqslant 0} \|\sigma(t, \cdot)\|_{C_b^2(R^N)} \vee \|b(t, \cdot)\|_{C_b^2(R^N)}$, such that:

$$\| \bar{Y}(T)\|_{L^p(\mathscr{W})}^p \leqslant A_p \int_0^T \| \bar{Y}(t)\|_{L^p(\mathscr{W})}^p\, dt + B_p T^{2p}, \qquad 0 \leqslant T \leqslant 1.$$

Clearly (5.11) follows from this plus Gromwall's inequality.

To prove (5.12), we start from

$$\mathscr{L}(x(T)) = \int_0^T [\mathscr{L}(\sigma(t, x(t))) - 1/2\sigma(t, x(t))]\, d\theta(t)$$
$$+ \int_0^T \mathscr{L}(b(t, x(t))\, dt, \qquad T \geqslant 0,$$

from which we see that there exist progressively measurable functions $\gamma_l^k(\cdot)$, $0 \leqslant k \leqslant d$ and $1 \leqslant l \leqslant N$, and $\beta^k(\cdot)$, $0 \leqslant k \leqslant d$, satisfying:

$$\max_{\substack{0 \leqslant k \leqslant d \\ 1 \leqslant l \leqslant N}} |\gamma_l^k(\cdot)| \leqslant K,$$

$$\max_{1 \leqslant k \leqslant d} |\beta^k(\cdot)| \leqslant K(1 + \operatorname{Trace}(A(\cdot))),$$

and

$$|\beta^0(\cdot)| \leqslant K \operatorname{Trace}(A(\cdot)),$$

where $K < \infty$ depends only on $\sup_{t \geqslant 0} \|\sigma(t, \cdot)\|_{C_b^2(R^N)} \vee \|b(t, \cdot)\|_{C_b^2(R^N)}$, such that:

$$\mathscr{L}(x_l(T)) = \sum_{k=1}^d \int_0^T \left(\sum_{l'=1}^N \gamma_{l'}^k(t)\, \mathscr{L}(x_{l'}(t)) + \beta^k(t) \right) d\theta_k(t)$$
$$+ \int_0^T \left(\sum_{l'=1}^N \gamma_{l'}^0(t)\, \mathscr{L}(x_{l'}(t)) + \beta^0(t) \right) dt, \qquad T \geqslant 0.$$

Proceeding as in the preceding paragraph, we derive from this that for each $p \in [2, \infty)$ there exist $A_p < \infty$ and $B_p < \infty$, with the same dependence as K, such that

$$\| \mathscr{L}x(T)\|_{L^p(\mathscr{W})}^p \leqslant A_p \int_0^T \| \mathscr{L}x(t)\|_{L^2(\mathscr{W})}^p\, dt + B_p T^{p/2}, \qquad 0 \leqslant T \leqslant 1.$$

Thus (5.12) is proved. Q.E.D.

(5.13) THEOREM. *Let $1 \leqslant M \leqslant N$ be given and set $A_{(M)}(t, x) = ((A_{m,n}(t, x)))_{1 \leqslant m,n \leqslant N}$ and $\Delta_{(M)}(t, x) = \det(A_{(M)}(t, x))$. Assume that for some $p \in [M, \infty)$, $0 < T \leqslant 1$, and $\nu > 0$:*

$$\| 1/\Delta_{(M)}(T, x)\|_{L^p(\mathscr{W})} \leqslant H_p T^{-\nu}. \tag{5.14}$$

Then the distribution $P_T^{(M)}(x, \cdot)$ of $x_{(M)}(T, x) = (x_1(T, x),..., x_M(T, x))$ admits a density $p_T^{(M)}(x, \cdot) \in \hat{C}(R^M)$ which is Hölder continuous with exponent $(1 - M/p)$ and constant depending only on p, H_p and

$\sup_{t\geqslant 0}\|\sigma(t,\ \cdot)\|_{C_b^2(R^N)} \vee \|b(t,\cdot)\|_{C_b^2(R^N)}$. *Moreover, there is a* $C<\infty$ *depending only on* p, H_p, *and* $\sup_{t\geqslant 0}\|\sigma(t,\ \cdot)\|_{C_b^2(R^N)} \vee \|b(t,\ \cdot)\|_{C_b^2(R^N)}$ *such that*

$$\| p_T(x,\ \cdot)\|_u \leqslant CT^{-\mu}, \tag{5.15}$$

where

$$\mu = M[(\nu+1/2-M)\vee 2(\nu-M)\vee 0]. \tag{5.16}$$

Proof. Throughout the proof we will neglect to mention the initial point $x\in R^N$.

By (1.15), if $F\in C_0^\infty(R^M)$, then:

$$E^{\mathscr{W}}\left[\frac{\partial F}{\partial x_m}(x_{(M)}(T))\right] = E^{\mathscr{W}}[F(x_{(M)}(T))\,\Psi_m((T))], \qquad 1\leqslant m\leqslant M,$$

with

$$\Psi_m = \sum_{n=1}^{M} (\langle x_n(T), A_{(M)}^{(m,n)}(T)/\Delta_{(M)}(T)\rangle + 2\mathscr{L}(x_m(T))\,A_{(M)}^{(m,n)}(t)/\Delta_{(M)}(T)),$$

where $A_{(M)}^{(m,n)}(T)$ denotes the (m,n)th-cofactor of $A_{(M)}(T)$. Set $q=(M+p)/2$. We want to estimate $\|\Psi_m(T)\|_{L^q(\mathscr{W})}$.

Take $\alpha = p/q$ and define β by $1/\beta = 1/2(1-1/\alpha)$. Then, by Hölder's inequality:

$$\begin{aligned}
&\|\mathscr{L}(x_n(T))\,A_{(M)}^{(m,n)}(T)/\Delta_{(M)}(T)\|_{L^q(\mathscr{W})}\\
&\quad\leqslant \|1/\Delta_{(M)}(T)\|_{L^p(\mathscr{W})}\,\|\mathscr{L}(x_n(T))\|_{L^{\beta q}(\mathscr{W})}\,\|A_{(M)}^{(m,n)}(T)\|_{L^{\beta q}(\mathscr{W})}\\
&\quad\leqslant M!H_p C_{\beta q} C_{(M-1)\beta q}^{M-1} T^{-\nu} T^{1/2} T^{M-1}\\
&\quad= CT^{-(\nu+1/2-M)}.
\end{aligned}$$

In deriving the second to last line, we have used (5.11) and (5.12).

Next, observe that

$$\begin{aligned}
\langle x_n(T), A_{(M)}^{(m,n)}(T)/\Delta_{(M)}(T)\rangle &= \langle x_n(T), A_{(M)}^{(m,n)}(T)\rangle/\Delta_{(M)}(T)\\
&\quad - \langle x_n(T), \Delta_{(M)}(T)\rangle\, A_{(M)}^{(m,n)}(T)/\Delta_{(M)}(T)^2.
\end{aligned}$$

Since $\langle x_n(T), A_{(M)}^{(m,n)}(T)\rangle/\Delta_{(M)}(T)$ is a finite linear combination of terms having the form

$$\langle x_n(T), A_{\mu,\nu}(T)\rangle\left(\prod_{i=1}^{M-2} A_{\mu_i,\nu_i}(T)\right)\Big/\Delta_{(M)}(T),$$

we can apply the technique just used above to obtain:

$$\|\langle x_n(T), A^{(m,n)}_{(M)}(T)\rangle/\Delta_{(M)}(T)\|_{L^q(\mathscr{W})} \leqslant CT^2T^{M-2}T^{-\nu}$$
$$\leqslant CT^{-2(\nu-M)}.$$

where $C<\infty$ has the desired dependence. By essentially the same argument:

$$\|\langle x_n(T), \Delta_{(M)}(T)\rangle\, A^{(m,n)}_{(M)}(T)/\Delta_{(M)}(T)^2\|_{L^q(\mathscr{W})} \leqslant CT^{-2(\nu-M)}.$$

Combining the preceding estimates, we arrive at

$$\|\Psi_m\|_{L^q(\mathscr{W})} \leqslant CT^{-\mu/M}.$$

Thus, if $\psi_m: R^M \to R^1$ is defined so that $\psi_m(x_{(M)}(T)) = E^{\mathscr{W}}[\Psi_m \mid \sigma(x_{(M)}(T))]$ (a.s., $\mathscr{W}$), then

$$\|\psi_m\|_{L^q(P_T^{(M)}(x,\cdot))} \leqslant \|\Psi_m\|_{L^q(\mathscr{W})} \leqslant CT^{-\mu/M}, \qquad 0<T\leqslant 1.$$

At the same time

$$\int_{R^M} \frac{\partial F}{\partial x_m}(y)\, P_T^{(M)}(x,dy) = \int_{R^M} F(y)\,\psi_m(y)\, P_T^{(M)}(x,dy)$$

for all $F \in C_0^\infty(R^M)$. Hence, by Lemma (1.18), the theorem has been proved.
Q.E.D.

(5.17) LEMMA. *Let $U_k(\cdot)$, $1\leqslant k\leqslant d$, $V(\cdot)$, and $X(\cdot)$ be as in Lemma (5.3). Then there exist constants $K<\infty$, $\delta>0$, and $\lambda>0$, depending only on the bounds on the $U_k(\cdot)$'s and $V(\cdot)$, such that:*

$$\mathscr{W}\big(\sup_{T_1\leqslant s\leqslant T_2} \|X(s,T_2)-I\|_{op} \geqslant 2\rho/(1-\rho)\big) \leqslant Ke^{-\lambda\rho^2/(T_2-T_1)}$$

for $\rho\in(0,1)$ and $0\leqslant T_1<T_2\leqslant T_1+\delta\rho$.

Proof. Set $X_{T_1}(\cdot)=X(T_1,\ \cdot)$. Then $X(s,T_2)=X_{T_1}(T_2)\,X_{T_1}(s)^{-1}$ and

$$X_{T_1}(T) = I + \sum_{k=1}^d \int_{T_1}^T U_k(t)\,X_{T_1}(t)\,d\theta_k(t) + \int_{T_1}^T V(t)\,X_{T_1}(t)\,dt,$$
$$T\geqslant T_1.$$

Thus, without loss of generality, we will assume that $T_1=0$.

Next, set $Y(t)=I-X(t)$. Assuming that $\|Y(s)\|_{op} \vee \|Y(t)\|_{op} \leqslant \rho<1$, we have:

$$X(s)^{-1} = I + \sum_1^\infty (Y(s))^n$$

and

$$X(s, t) = (I - Y(t)) \left(I + \sum_1^\infty (Y(s))^n\right)$$

$$= I - Y(t) + \sum_1^\infty (Y(s))^n - Y(t) \sum_1^\infty (Y(s))^n,$$

and so in particular:

$$\|I - X(s, t)\|_{op} \leqslant \rho + \frac{\rho}{1-\rho} + \frac{\rho^2}{1-\rho} = 2\rho/(1-\rho).$$

Thus:

$$\mathscr{W}(\sup_{0 \leqslant s \leqslant t} \|X(s, t) - I\|_{op} \geqslant 2\rho/(1-\rho)) \leqslant \mathscr{W}(\sup_{0 \leqslant s \leqslant t} \|X(s) - I\|_{op} \geqslant \rho)$$

$$\leqslant \mathscr{W}(\sup_{0 \leqslant s \leqslant t} \|X(s) - I\|_{H.S.} \geqslant \rho).$$

To complete the proof, set $\tau = \inf\{t \geqslant 0: \|X(t) - I\|_{H.S.} \geqslant \rho\}$. Clearly, $\mathscr{W}(\sup_{0 \leqslant s \leqslant t} \|X(s) - I\|_{H.S.} \geqslant \rho) = \mathscr{W}(\|X(t \wedge \tau) - I\|_{H.S.} \geqslant \rho)$. Note that

$$X(t \wedge \tau) = I + \sum_{k=1}^d \int_0^t \tilde{U}_k(s)\, d\theta_k(s) + \int_0^t \tilde{V}(s)\, ds,$$

where the $\tilde{U}_k(\cdot)$'s and $\tilde{V}(\cdot)$'s are bounded in terms of the $U_k(\cdot)$'s and $V(\cdot)$'s, respectively. Hence if $\delta > 0$ is chosen so that $\delta \sup_s \|\tilde{V}(s)\|_{H.S.} = 1/2$, then for $t \leqslant \delta\rho$:

$$\mathscr{W}(\|X(t \vee \tau) - I\|_{H.S.} \geqslant \rho) \leqslant \mathscr{W}\left(\left\| \sum_{k=1}^d \int_0^t \tilde{U}_k(s)\, d\theta_k(s) \right\|_{H.S.} \geqslant \rho/2\right)$$

$$\leqslant K \exp(-\lambda\rho^2/t). \qquad \text{Q.E.D.}$$

(5.18) THEOREM. *Let $1 \leqslant M \leqslant N$ and assume that $a_{(M)}(\cdot) \geqslant \varepsilon I$ for some $\varepsilon > 0$, where $a_{(M)}(t, \cdot) = ((a_{m,n}(t, \cdot))))_{1 \leqslant m,n \leqslant M}$. Then for $0 < T \leqslant 1$, the distribution $P_T^{(M)}(x, \cdot)$ of $x_{(M)}(T) \equiv (x_1(T),..., x_M(T))$ under $\mathscr{W}$ admits a density $p_T^{(M)}(x, \cdot) \in \hat{C}(R^M)$ which is Hölder continuous with every exponent strictly less than 1. Furthermore, there exists a $C < \infty$ depending only on ε and $\sup_{t \geqslant 0} \|\sigma(t, \cdot)\|_{C_b^2(R^N)} \vee \|b(t, \cdot)\|_{C_b^2(R^N)}$ such that $\|p_T^{(M)}(x, \cdot)\|_u \leqslant CT^{-M/2}$, $0 < T \leqslant 1$. Finally, for fixed $0 < T \leqslant 1$ and $\alpha \in (0, 1)$, the α-Hölder constant of $p_T^{(M)}(x, \cdot)$, is dominated by a quantity depending only on ε and $\sup_{t \geqslant 0} \|\sigma(t, \cdot)\|_{C_b^2(R^N)} \vee \|b(t, \cdot)\|_{C_b^2(R^N)}$.*

Proof. In view of Theorem (5.13), all that we need to prove is that (in the notation of Theorem (5.13)): $\| 1/\Delta_{(M)}(T)\|_{L^p(\mathscr{W})} \leqslant H_p T^{-M}$, $0 < T \leqslant 1$ and $p \in [1, \infty)$, where $H_p < \infty$ depends only on ε and $\sup_{t \geqslant 0} \|\sigma(t, \cdot)\|_{C_b^3(R^N)} \vee \|b(t, \cdot)\|_{C_b^3(R^N)}$.

Let

$$\Lambda = \sup_{\substack{t \geqslant 0 \\ x \in R^N}} \|a(t, x)\|_{op}$$

and choose $\rho \in (0, 1/2)$ so that $\Lambda 4\rho/(1-\rho)^2 < \varepsilon/2$. Given $0 < T \leqslant 1$, define $\zeta = \sup\{\delta \in [0, T]: \|X(s, T) - I\|_{op} \leqslant 2\rho/(1-\rho)$ for all $s \in [T-\delta, T]\}$. By Lemma 5.17), there exist $K < \infty$, $\delta_0 > 0$, and $\lambda > 0$, depending only on $\sup_{t \geqslant 0} \|\sigma(t, \cdot)\|_{C_b^1(R^N)} \vee \|b(t, \cdot)\|_{C_b^1(R^N)}$, such that for $\delta \leqslant T \wedge (\delta_0 \rho)$: $\mathscr{W}(\zeta \leqslant \delta) \leqslant K \exp(-\lambda \rho^2/\delta)$. Hence $\| 1/\zeta^M\|_{L^p(\mathscr{W})} \leqslant K_p T^{-M}$, $0 < T \leqslant 1$ and $p \in [1, \infty)$, where K_p depends only on $\sup_{t \geqslant 0} \|\sigma(t, \cdot)\|_{C_b^1(R^N)} \vee \|b(t, \cdot)\|_{C_b^1(R^N)}$. At the same time:

$$\begin{aligned} A_{(M)}(T) &= \int_0^T (X(s, T)\, a(s, x(s))\, X(s, T)^*)_{(M)}\, ds \\ &\geqslant \int_{T-\zeta}^T (X(s, T)\, a(s, x(s))\, X(s, T)^*)_{(M)}\, ds \\ &\geqslant (\varepsilon - (4\rho/(1-\rho)^2)\Lambda)\, \zeta I_{(M)} \\ &\geqslant \varepsilon/2 \zeta I_{(M)}. \end{aligned}$$

Hence the required estimate on $\| 1/\Delta_{(M)}(T)\|_{L^p(\mathscr{W})}$ follows. Q.E.D.

(5.19) *Remarks.* Two technical comments about the hypotheses in Theorem (5.18) are in order. In the first place, the restriction that T be less than 1 is clearly inessential: the same conclusions hold for $T \in (0, T_0]$ for every $T_0 < \infty$. Secondly, we do not need three spacial derivatives of σ and b. Indeed, since all of our estimates depend only on two derivatives, two derivatives suffice. The proof is immediate from what we already have plus an easy limit argument.

(5.20) *Remark.* It is gratifying that our estimate on $\| p_T^{(M)}(x, \cdot)\|_u$ is in terms on $T^{-M/2}$, which is precisely what one would suspect to be the correct behavior of $\| p_T^{(M)}(x, \cdot)\|_u$ near 0. One should also expect that $p_T^{(M)}(x, y)_{(M)}$ satisfies an appropriate lower bound. Such an estimate is well-known from the classical parametrix method when $M = N$ (cf. [2]). In the present case, the best that I have been able to show is that $p_T^{(M)}(x, y_{(M)}) > 0$ for $y_{(M)}$ in a dense open subset of R^M. This last observation is an easy application of the results in [13].

(5.21) *Remark.* By combining the ideas used in this section with those developed by Watenabe (cf. [6]) one can prove the following variant of Theorem (5.18). Let L be given in Hörmander's form; that is, $L=\sum_1^r X_j^2+X_0$, where $X_0,\dots,X_r$ are $C_b^\infty(R^N)$ vector fields on R^N. Assume that $\partial/\partial x_1|_x,\dots,\partial/\partial x_M|_x\in Lie(X_1,\dots,X_r)(x)$ for all $x\in R^N$ and let $P(T,x,\ \cdot)$ denote the transition probability function for the diffusion generated by L. Then if $P^{(M)}(T,x,\ \cdot)$ denotes the marginal distribution under $P^{(M)}(T,x,\ \cdot)$ of coordinates 1 through M, $P^{(M)}(T,x,\ \cdot)$ admits a density $p^{(M)}(T,x,\ \cdot)\in\hat{C}(R^M)$ and $p^{(M)}(T,x,\ \cdot)$ has the smoothness properties described in Theorem (5.18). Moreover, $\|p^{(M)}(T,x,\ \cdot)\|\leqslant CT^{-\nu}$, $0<T\leqslant 1$, where $\nu>0$ can be estimated in terms of M and the number of Lie operations required to generate $\partial/\partial x_1,\dots,\partial/\partial x_M$ from $X_1,\dots,X_r$.

6. Higher Order Derivatives

We continue with the notation described at the beginning of Section 5. Our goal is to produce a sizeable class $\mathscr{G}\subseteq\mathscr{K}(\mathscr{L})$ with the following closure properties:

$$\begin{aligned}&\text{(i) if } N\geqslant 1,\ F\in C_{\nearrow}^\infty(R^N), \text{ and } \Phi_1,\dots,\Phi_N\in\mathscr{G}, \text{ then } F\circ\Phi\in\mathscr{G},\\&\text{(ii) if } \Phi\in\mathscr{G}, \text{ then } \mathscr{L}(\Phi)\in\mathscr{G}.\end{aligned}\tag{6.1}$$

Clearly, (i) and (ii) imply that for all $\Phi,\Psi\in\mathscr{G}$: $\langle\Phi,\Psi\rangle\in\mathscr{G}$.

Let $D\geqslant 1$ and V: $[0,\infty)\times R^D\to R^D$ be given. We will say that $V(\cdot)$ is *lower triangular with respect to the grading* $\{D_\mu\}_{\mu=0}^M$ if $V(\cdot)$ is measurable, $0=D_0<\cdots<D_M=D$, and

$$V(t,X)=\begin{pmatrix}V_{(1)}(t,X_{(1)})\\ \vdots\\ V_{(M)}(t,X_{(M)})\end{pmatrix}\in R^{d_1}\times\cdots\times R^{d_M},\qquad t\geqslant 0\quad\text{and}\quad X\in R^D,$$

where $X_{(\mu)}\equiv(X_1,\dots,X_{D_\mu})$, $d_\mu=D_\mu-D_{\mu-1}$, and $V_{(\mu)}(t,\cdot)\in(C^\infty(R^{D_\mu}))^{d_\mu}$, $t\geqslant 0$, has the property that for each $\alpha\in(\mathscr{N})^{D_\mu}$ there is a $C_\alpha<\infty$ and a $\gamma_\alpha\geqslant 0$ such that $\gamma_\alpha=0$ when $\max_{D_{\mu-1}<j\leqslant D_\mu}\alpha_j\geqslant 1$ and $\sup_{t\geqslant 0}|D^\alpha_{X_{(\mu)}}V_{(\mu)}(t,X_{(\mu)})|\leqslant C_\alpha(1+|X_{(\mu)}|)^{\gamma_\alpha}$, $X_{(\mu)}\in R^{D_\mu}$.

Given σ: $[0,\infty)\times R^D\to R^D\otimes R^d$ and b: $[0,\infty)\times R^D\to R^D$, we say that $(\sigma(\cdot),b(\cdot))$ is a *lower triangular system* of coefficients if the column vector fields $(\sigma_1^\bullet(\cdot),\dots,\sigma_d^\bullet(\cdot)$ and the vector field $b(\cdot)$ are simultaneously lower triangular with respect to some common grading.

(6.2) Theorem. *Let $(\sigma(\cdot),b(\cdot))$ be a lower triangular system of coef-*

ficients and let $\Xi_0 \in R^D$ *be given. Then there is a unique progressively measurable* Ξ: $[0, \infty) \times \Theta \to R^D$ *satisfying*

$$\Xi(T) = \Xi_0 + \int_0^T \sigma(t, \Xi(t))\, d\theta(t) + \int_0^T b(t, \Xi(t))\, dt, \qquad T \geqslant 0. \tag{6.3}$$

Furthermore, $\Xi(T) \in (\mathscr{K}(\mathscr{L}))^D$, $T \geqslant 0$, $T \to (\Xi(T), \langle\langle \Xi(T), \Xi(T) \rangle\rangle, \mathscr{L}(\Xi)(T)))$ *is continuous (a.s.,* $\mathscr{W}$*), and* $|||\Xi(\cdot)|||_{(p),T} < \infty$ *for all* $T \geqslant 0$ *and* $p \in [4, \infty)$. *Finally, there is a lower triangular system* $(\tilde{\sigma}(\cdot), \tilde{b}(\cdot))$ *and a* $\tilde{\Xi}_0 \in R^D \times R^{D^2} \times R^D$ *such that*

$$\tilde{\Xi}(T) = \tilde{\Xi}_0 + \int_0^T \tilde{\sigma}(t, \tilde{\Xi}(t))\, d\theta(t) + \int_0^T \tilde{b}(t, \tilde{\Xi}(t))\, dt, \qquad T \geqslant 0, \tag{6.4}$$

where

$$\tilde{\Xi}(\cdot) = \begin{pmatrix} \Xi(\cdot) \\ \langle\langle \Xi(\cdot), \Xi(\cdot) \rangle\rangle \\ \mathscr{L}(\Xi(\cdot)) \end{pmatrix}.$$

Proof. Using Theorem (4.11) and working by induction on M, we see that it is sufficient to check (under the assumptions that $\Xi(T) \in (\mathscr{K}(\mathscr{L}))^D$, $T \geqslant 0$, and that $|||\Xi(\cdot)|||_{(p),T} < \infty$, $T \geqslant 0$ and $p \in [4, \infty)$) that $\tilde{\Xi}(\cdot)$ is given by (6.4) with $(\tilde{\sigma}(\cdot), \tilde{b}(\cdot))$ being a lower triangular system. The checking of this fact entails a certain amount of bookkeeping.

Set $x^{(\mu)} = (X_{D_{\mu-1}+1}, \dots, X_{D_{\mu-1}+d_\mu})$, $1 \leqslant \mu \leqslant M$ and write

$$\sigma(t, X) = \begin{pmatrix} \sigma^{(1)}(t, x^{(1)}) \\ \vdots \\ \sigma^{(M)}(t, x^{(1)}, \dots, x^{(M)}) \end{pmatrix} \in (R^{d_1} \times \cdots \times R^{d_M}) \otimes R^d$$

and

$$b(t, X) = \begin{pmatrix} b^{(1)}(t, x^{(1)}) \\ \vdots \\ b^{(M)}(t, x^{(1)}, \dots, x^{(M)}) \end{pmatrix} \in R^{d_1} \times \cdots \times R^{d_M}$$

for $X \in R^D$. (We are, of course, assuming that $\{D_\mu\}_{\mu=0}^M$ is the grading for $(\sigma(\cdot), b(\cdot))$.) Using the convention that $(m', n') < (m, n)$ if $m' < m$ or if $m' = m$ and $n' < n$, set $y^{(\mu,\nu)} = (y^{(\mu,\nu)}_{(1,1)}, \dots, y^{(\mu,\nu)}_{(d_\mu, d_\nu)}) \in R^{d_\mu \cdot d_\nu}$ for $1 \leqslant \mu, \nu \leqslant M$

and let $Y_{(\mu,\nu)} = (y^{(1,1)},\ldots, y^{(\mu,\nu)}) \in R^{D_\mu \cdot D_\nu}$. Also, set $z^{(\mu)} = (z_1^{(\mu)},\ldots, z_{d_\mu}^{(\mu)}) \in R^{d_\mu}$ and $Z_{(\mu)} = (z^{(1)},\ldots, z^{(\mu)})$ for $1 \leqslant \mu \leqslant M$. Finally, define

$$\begin{aligned}
\tilde{\sigma}^{(\mu,\nu)}_{(i,j),k}(t; X; Y_{(\mu,\nu)}) = {} & \sum_{\mu'=1}^{\mu} \sum_{i'=1}^{d_{\mu'}} \frac{\partial \sigma^{(\mu)}_{i,k}}{\partial x^{(\mu')}_{i'}}(t, X) \cdot y^{(\mu',\nu)}_{(i',j)} \\
& + \sum_{\nu'=1}^{\nu} \sum_{j'=1}^{d_{\nu'}} \frac{\partial \sigma^{(\nu)}_{j,k}}{\partial x^{(\nu')}_{j'}}(t, X) \cdot y^{(\mu,\nu')}_{(i,j')},
\end{aligned}$$

$$\begin{aligned}
\tilde{b}^{(\mu,\nu)}_{(i,j)}(t, X; Y_{(\mu,\nu)}) = {} & \sum_{\mu'=1}^{\mu} \sum_{i'=1}^{d_{\mu'}} \frac{\partial b^{(\mu)}}{\partial x^{(\mu')}_{i'}}(t, X) \cdot y^{(\mu',\nu)}_{(i',j)} \\
& + \sum_{\nu'=1}^{\nu} \sum_{j'=1}^{d_{\nu'}} \frac{\partial b^{(\nu)}}{\partial x^{(\nu')}_{j'}}(t, X) \cdot y^{(\mu,\nu')}_{(i,j')} \\
& + \sum_{k=1}^{d} \sum_{\mu'=1}^{\mu} \sum_{\nu'=1}^{\nu} \sum_{i'=1}^{d_{\mu'}} \sum_{j'=1}^{d_{\nu'}} \frac{\partial \sigma^{(\mu)}_{i,k}}{\partial x^{(\mu')}_{i'}}(t, X) \\
& \times \frac{\partial \sigma^{(\nu)}_{j,k}}{\partial x^{(\nu')}_{j'}}(t, X) \cdot y^{(\mu',\nu')}_{(i',j')},
\end{aligned}$$

$$\begin{aligned}
\tilde{\sigma}^{(\mu)}_{i,k}(t; X; Y; Z_{(\mu)}) = {} & \sum_{\mu'=1}^{\mu} \sum_{i'=1}^{d_{\mu'}} \frac{\partial \sigma^{(\mu)}_{i,k}}{\partial x^{(\mu')}_{i'}}(t, X) \cdot z^{(\mu')}_{i'} \\
& + 1/2 \sum_{\mu',\mu''=1}^{\mu} \sum_{i'=1}^{d_{\mu'}} \sum_{i''=1}^{d_{\mu''}} \frac{\partial^2 \sigma^{(\mu)}_{i,k}}{\partial x^{(\mu')}_{i'} \partial x^{(\mu'')}_{i''}}(t, X) \\
& \cdot y^{(\mu',\mu'')}_{(i',i'')} - 1/2\sigma^{(\mu)}_{i,k}(t, X),
\end{aligned}$$

and

$$\begin{aligned}
\tilde{b}^{(\mu)}_{i}(t; X; Y; Z_{(\mu)}) = {} & \sum_{\mu'=1}^{\mu} \sum_{i'=1}^{d_{\mu'}} \frac{\partial b^{(\mu)}_{i}}{\partial x^{(\mu')}_{i'}}(t, X) \cdot z^{(\mu')}_{i'} \\
& + 1/2 \sum_{\mu',\mu''=1}^{\mu} \sum_{i'=1}^{d_{\mu'}} \sum_{i''=1}^{d_{\mu''}} \frac{\partial^2 b^{(\mu)}}{\partial x^{(\mu')}_{i'} \partial x^{(\mu'')}_{i''}}(t, X) \\
& \times y^{(\mu',\mu'')}_{(i',i'')}.
\end{aligned}$$

where $Y \equiv Y_{(M,M)}$. Then the system $(\tilde{\sigma}(\cdot), \tilde{b}(\cdot))$ given by

$$\tilde{\sigma}(t; X; Y; Z) = \begin{bmatrix} \sigma^{(1)}(X_{(1)}) \\ \vdots \\ \sigma^{(M)}(X_{(M)}) \\ \tilde{\sigma}^{(1,1)}(X; Y_{(1,1)}) \\ \vdots \\ \tilde{\sigma}^{(M,M)}(X; Y_{(M,M)}) \\ \tilde{\sigma}^{(1)}(X; Y; Z_{(1)}) \\ \vdots \\ \tilde{\sigma}^{(M)}(X; Y; Z_{(M)}) \end{bmatrix}$$

and

$$\tilde{b}(t; X; Y; Z) = \begin{bmatrix} b^{(1)}(X_{(1)}) \\ \vdots \\ b^{(M)}(X_{(M)}) \\ \tilde{b}^{(1,1)}(X; Y_{(1,1)}) \\ \vdots \\ \tilde{b}^{(M,M)}(X; Y_{(M,M)}) \\ \tilde{b}^{(1)}(X; Y; Z_{(1)}) \\ \vdots \\ \tilde{b}^{(M)}(X; Y; Z_{(M)}) \end{bmatrix},$$

where $Z = Z_{(M)}$, is lower triangular. Moreover, using (4.9) and (4.7), one can easily check hat $\tilde{\Xi}(\cdot)$ satisfies (6.3) with this choice of $(\tilde{\sigma}(\cdot), \tilde{b}(\cdot))$ and

$$\tilde{\Xi} = \begin{pmatrix} \Xi_0 \\ 0 \\ 0 \end{pmatrix}. \qquad \text{Q.E.D.}$$

We now define $\mathscr{S}_0$ to be the set of all $\Phi: \Theta \to R^1$ such that there exist $D \geqslant 1$, $\Xi_0 \in R^D$, a lower triangular system $(\sigma(\cdot), b(\cdot))$, and $T \geqslant 0$ with the property that $\Phi = \Xi_D(T)$, where $\Xi(\cdot)$ is the progressively measurable solution to (6.3). Using Itô's formula, one can easily check that $\mathscr{S}$ posseses property (i) of (6.1); and using Theorem (6.2), one sees that $\mathscr{S}$ has property (ii) of (6.1). Let $\mathscr{S}$ be the set of Φ for which there exists a sequence $\{\Phi_n\}_1^\infty \subseteq \mathscr{S}_0$ such that $\Phi_n \to \Phi$ in $L^q(\mathscr{W})$ and $\lim_{m\to\infty} \sup_{n\geqslant m} \|\mathscr{L}^k(\Phi_n) - \mathscr{L}^k(\Phi_m)\|_{L^q(\mathscr{W})} = 0$ for every $q \in (2, \infty)$ and $k \geqslant 1$. Note that $\mathscr{S}$ again has properties (i) and (ii) of (6.1). Moreover, for $\Phi \in \mathscr{S}_0^+$ with $1/\Phi \in \bigcap_1^\infty L^q(\mathscr{W})$, $1/\Phi \in \mathscr{S}$.

(6.5) THEOREM. *Suppose that* $\Phi = (\Phi_1, ..., \Phi_N) \in (\mathscr{S})^N$ *and set* $\Delta = \det(\langle\langle \Phi, \Phi \rangle\rangle)$ *and* $\mu = \mathscr{W} \circ \Phi^{-1}$. If $1/\Delta \in \bigcap_1^\infty L^p(\mathscr{W})$, *then* μ *admits a density* $f \in \hat{C}^\infty(R^N)$.

Proof. Let $\mathscr{H}_k$, $1 \leqslant k \leqslant N$, be defined as in (1.16) relative to Φ; and define $\bar{\mathscr{H}}_k \Psi = \mathscr{H}_k(\Psi/\Delta)$ for $\Psi \in \mathscr{K}(\mathscr{L})$. Given an $\alpha \in (\mathscr{N})^N$, set $\bar{\mathscr{H}}^{(\alpha)} = (\bar{\mathscr{H}}_1)^{\alpha_1} \circ \cdots \circ (\bar{\mathscr{H}}_N)^{\alpha_N}$. Using induction, we see that for any $F \in C_0^\infty(R^N)$:

$$\int_{R^N} D^\alpha F \, d\mu = (-1)^{|\alpha|} E^{\mathscr{W}}[(F \circ \Phi) \, \Psi^{(\alpha)}],$$

where $\Psi^{(\alpha)} \equiv \bar{\mathscr{H}}^{(\alpha)}(1) \in \mathscr{K}(\mathscr{L})$. Hence we obtain

$$\int_{R^N} D^\alpha F \, d\mu = (-1)^{|\alpha|} \int_{R^N} F \psi^{(\alpha)} \, d\mu, \qquad F \in C_{0(R}^{\infty}{}^{N)}, \tag{6.6}$$

where $\psi^{(\alpha)}: R^N \to R^1$ satisfies: $\psi^{(\alpha)} \circ \Phi = E^{\mathscr{W}}[\Psi^{(\alpha)} \mid \sigma(\Phi^{-1}(\mathscr{B}_{R^N}))]$. In particular, $\|\psi^{(\alpha)}\|_{L^q(\mu)} \leqslant \|\Psi^{(\alpha)}\|_{L^q(\mathscr{W})} < \infty$ for all $\alpha \in \mathscr{N}^N$ and $1 \leqslant q < \infty$.

From Lemma (1.18), (6.6) with $|\alpha| = 1$ already implies that μ admits a density $f \in \hat{C}(R^N)$. Moreover, (6.6) is equivalent to the statement $D^\alpha f = \psi^{(\alpha)} f$ in the sense of distributions. But $\|\psi^{(\alpha)} f\|_{L^q(R^N)} \leqslant \|f\|_u^{1-1/p} \|\psi^{(\alpha)}\|_{L^q(\mu)} < \infty$. Thus $D^\alpha f \in L^q(R^N)$ for all $\alpha \in (\mathcal{N})^N$ and $1 \leqslant q < \infty$. In particular, the standard Sobolev embedding theorem implies that $f \in \hat{C}^\infty(R^N)$. Q.E.D.

7. Some Applications of Higher Order Derivatives

Again we work with the notation introduced at the beginning of Section 5. Let $\sigma: R^N \to R^N \otimes R^N$ and $b: R^N \to R^N$ be C^∞-functions satisfying $\|D^\alpha \sigma\|_u \vee \|D^\alpha b\|_u < \infty$ for all $|\alpha| \geqslant 1$ (cf. Remark (7.8) below). Given $x \in R^N$, let $x(\cdot, x)$ denote the solution to

$$x(T, x) = x + \int_0^T \sigma(x(t, x))\, d\theta(t) + \int_0^T b(x(t, x))\, dt, \qquad T \geqslant 0. \tag{7.1}$$

Then the distribution $P(T, x, \cdot)$ of $x(T, x)$ under $\mathscr{W}$ is the transition probability function for the diffusion whose quasi-generator is

$$L = 1/2 \sum_{i,j=1}^N a^{ij}(x) \frac{\partial^2}{\partial x_i\, \partial x_j} + \sum_{i=1}^N b^i(x) \frac{\partial}{\partial x_i}, \tag{7.2}$$

where $a(\cdot) \equiv \sigma\sigma^*(\cdot)$. Thus far we have been using he Malliavin calculus to study $P(T, x, \cdot)$ as a function of the "forward variable." We are now going to refine the results already obtained about the "forward variable" as well as show how the Malliavin calculus can be used to study the "backward variable." However, before we can discuss the "backward variable," it will be necessary to review some facts about $x(T, x)$ as a function of $x \in R^N$. Most of what we need to recall are facts that have been more or less well-known for some time (cf. [3]). Recently several authors have provided more rigorous treatments than the original authors. For a treatment that is closest to what is required here, the reader may want to look at the exposition of the ideas of Kunita [7] as presented in [15].

The central fact which we need is the existence of a $\xi: [0, \infty) \times R^N \times \Theta \to R^N$ such that: $\xi(\cdot, x)$ is progressively measurable for each $x \in R^N$; $\xi(T, \cdot) \in C^\infty(R^N)$ for all $T \geqslant 0$ (a.s., $\mathscr{W}$); $(T, x) \to D_x^\alpha \xi(T, x)$ is continuous (a.s., $\mathscr{W}$) for each $\alpha \in \mathcal{N}^N$; and $\xi(\cdot, x) = x(\cdot, x)$ (a.s., $\mathscr{W}$) for each $x \in R^N$. An essential ingredient in the proof of the existence of $\xi(\cdot)$ is the observation that for every $m \geqslant 0$ the vector $\Xi_{(m)}(\cdot, x) \equiv \{D^\alpha \xi(\cdot, x): |\alpha| \leqslant m\}$ solves a stochastic integral equation in which the coefficients are a lower triangular system when the entries of $\Xi_{(m)}(\cdot, x)$ are ordered so that $D^\alpha \xi(\cdot, x)$ precedes $D^\beta \xi(\cdot, x)$ if $|\alpha| < |\beta|$ and the grading is taken according to $|\alpha|$. In particular, $D_x^\alpha \xi(T,x) \in \mathscr{S}$ for $\alpha \in \mathscr{N}^N$, $T \geqslant 0$, and $x \in R^N$.

(7.3) LEMMA. *Given $F \in C_0^\infty(R^N)$, set $u_F(T,x) = \int F(y)\,P(T,x,dy)$, $(T,x) \in [0,\infty) \times R^N$. Then $u_F \in C^\infty([0,\infty) \times R^N)$. Moreover, for each $\alpha \in (\mathscr{N})^N$ there exist polynomials $\{P_{\alpha,\beta} : \beta \leqslant \alpha\}$ such that:*

$$(D_x^\alpha u_F)(T,x) = \sum_{\beta \leqslant \alpha} E^{\mathscr{W}}[(D^\beta F)(\xi(T,x))\,P_{\alpha,\beta}(\Xi_{(|\alpha|)}(T,x))] \tag{7.4}$$

for all $T \geqslant 0$, $x \in R^N$, and $F \in C_0^\infty(R^N)$.

Proof. The proof of (7.4) is a simple application of the chain rule. Once one has (7.4), the differentiability of u_F with respect to T follows from $\partial u_F/\partial T = u_F$, $T \geqslant 0$. Q.E.D.

Lemma (7.3) simply says that, so long as $\sigma(\cdot)$ and $b(\cdot)$ are smooth, smooth initial data are taken into smooth functions by the diffusion semigroup determined by L. Obviously, one can expect no better statement without imposing non-degeneracy assumptions on L. The next theorem provides an example of the sort of statement that one can prove using the Malliavin calculus when sufficient non-degeneracy is assumed.

(7.4) THEOREM. *For $1 \leqslant M \leqslant N$, set $\xi_{(M)}(\cdot,x) = (\xi_1(\cdot,x),\dots,\xi_M(\cdot,x))$ and $\Delta_{(M)}(\cdot,x) = \det(\langle\langle \xi_{(M)}(\cdot,x), \xi_{(M)}(\cdot,x)\rangle\rangle)$. If, for some $T > 0$ and $x \in R^N$, $1/\Delta_{(M)}(T,x) \in \bigcap_1^\infty L^p(\mathscr{W})$, then for every $k \geqslant 0$ there is a $C_k < \infty$ such that*

$$\max_{|\alpha| \leqslant k} |D_x^\alpha u_F(T,x)| \leqslant C_k \max\{\|D^\beta F\|_u : |\beta| \leqslant k \text{ and } \beta_1 = \cdots = \beta_M = 0\}, \tag{7.5}$$

for all $F \in C_0^\infty(R^N)$, where $u_F(T,x) \equiv \int F(y)\,P(T,x,dy)$. Moreover, if $1/\Delta_{(M)}(T,x) \in \bigcap_1^\infty L^p(\mathscr{W})$ for all $(T,x) \in (T_1,T_2) \times U$, where $0 \leqslant T_1 < T_2$ and $U \subseteq R^N$ is bounded and open, and if

$$\sup_{\substack{T_1 < T < T_2 \\ x \in U}} \|1/\Delta_{(M)}(T,x)\|_{L^p(\mathscr{W})} < \infty$$

for all $p \in [1,\infty)$, then the C_k in (7.5) can be chosen independent of $(T,x) \in (T_1,T_2) \times U$. In particular, if $F \in B(R^N)$ satisfies $F(x_{(M)},\cdot) \in C_b^\infty(R^{N-M})$ for all $x_{(M)} \equiv (x_1,\dots,X_M) \in R^M$ and $\sup_{x_{(M)} \in R^M} \|F(x_{(M)},\cdot)\|_{C_b^k(R^{N-M})} < \infty$ for all $k \geqslant 0$, then $u_F \in C_b^\infty((T_1,T_2) \times U)$ and $\partial u_F/\partial T = Lu_F$ in $(T_1,T_2) \times U$.

Proof. Set $A(T,x) = \langle\langle \xi(T,x), \xi(T,x)\rangle\rangle$ and let $A_{(M)}^{(k,l)}(T,x)$ denote the (k,l)th cofactor of $\langle\langle \xi_{(M)}(T,x), \xi_{(M)}(T,x)\rangle\rangle$. Given $F \in C_0^\infty(R^N)$, we have

$$\langle F(\xi(T,x)), \xi_l(T,x)\rangle = \sum_{j=1}^N \frac{\partial F}{\partial x_j}(\xi(T,x))\,A_{jl}(T,x).$$

Hence, for $1 \leqslant k \leqslant M$:

$$\sum_{l=1}^{M} \langle F(\xi(T,x)), \xi_l(T,x)\rangle A_{(M)}^{(k,l)}(T,x) = \Delta_{(M)}(T,x) \frac{\partial F}{\partial x_k}(\xi(T,x)) + \sum_{j=M+1}^{N} \left(\sum_{l=1}^{M} A_{(M)}^{(k,l)}(T,x) A_{jl}(T,x)\right) \frac{\partial F}{\partial x_j}(\xi(T,x)).$$

Proceeding as in the proof of Theorem (1.14), we now see that for $1 \leqslant k \leqslant M$ and $\Psi \in \mathscr{S}$:

$$E^{\mathscr{W}}\left[\frac{\partial F}{\partial x_k}(\xi(T,x))\Psi\right] = E^{\mathscr{W}}[F(\xi(T,x))\mathscr{R}_0\Psi] + \sum_{j=M+1}^{N} E^{\mathscr{W}}\left[\frac{\partial F}{\partial x_j}(\xi(T,x))\mathscr{R}_j\Psi\right], \quad (7.6)$$

where $\mathscr{R}_0$ and $\mathscr{R}_{M+1},\ldots,\mathscr{R}_N$ map $\mathscr{S}$ linearly into $\mathscr{S}$.

Starting with (7.4) and making repeated applications of (7.6), we arrive at (7.5) by a simple induction argument.

The remainder of the theorem is an easy consequence of (7.5) once one has noticed that the stated assumptions provide uniform estimates on the quantities appearing on the right hand sides of (7.4) and (7.6). Q.E.D.

Another application of the same sort is the following theorem about $P(T,x,\cdot)$.

(7.7) THEOREM. *Suppose that* $1/\Delta_{(M)}(T,x) \in \bigcap_1^\infty L^p(\mathscr{W})$ *for all* $(T,x) \in (T_1,T_2) \times U$, *where* $0 \leqslant T_1 < T_2$ *and* U *is an open set in* R^N. *Assume that for each* $p \in [1,\infty)$ $\|1/\Delta_{(M)}(T,x)\|_{L^p(\mathscr{W})}$ *is uniformly bounded for* (T,x) *in compact subsets of* $(T_1,T_2) \times U$. *Then for each* $(T,x) \in (T_1,T_2) \times U$ *the distribution* $P^{(M)}(T,x,\cdot)$ *of* $\xi_{(M)}(T,x)$ *under* $\mathscr{W}$ *admits a density* $p^{(M)}(T,x,\cdot) \in \hat{C}(R^M)$ *and* $p^{(M)} \in C^\infty((T_1,T_2) \times U \times R^M)$.

Proof. Using (7.6), we see that for any $\alpha \in (\mathscr{N})^N$ and $F \in C_0^\infty(R^M)$:

$$D_x^\alpha \int_{R^M} F(y)\, P^{(M)}(T,x,dy) = E^{\mathscr{W}}[F(\xi(T,x))\,\Psi_\alpha(T,x)],$$

where $(T,x) \in (T_1,T_2) \times U \to \Psi_\alpha(T,x) \in \mathscr{S}$ is continuous (a.s., $\mathscr{W}$). Replacing F by $D_y^\beta F$, $\beta \in (\mathscr{N})^M$, and proceeding as in Theorem (1.14), we arrive at

$$D_x^\alpha \int_{R^M} (D^\beta F)(y)\, P^{(M)}(T,x,dy) = (-1)^{|\beta|} E^{\mathscr{W}}[F(\xi(T,x))\,\Psi_{\alpha;\beta}(T,x)],$$

where $(T, x) \in (T_1, T_2) \times U \to \Psi_{\alpha,\beta}(T, x) \in \mathscr{G}$ is continuous (a.s. $\mathscr{W}$). Thus

$$D_x^\alpha \int_{R^M} (D^\beta F)(y)\, P^{(M)}(T, x, dy) = (-1)^{|\beta|} \int_{R^M} F(y)\, \psi_{\alpha,\beta}(T, x, y)\, P^{(M)}(T, x, dy)$$

for all $\alpha \in (\mathscr{N})^N$, $\beta \in (\mathscr{N})^M$, and $(T, x) \in (T_1, T_2) \times U$, where $\psi_{\alpha,\beta}(T, x, \cdot) \in \bigcap_1^\infty L^q(P^{(M)}(T, x, \cdot))$. In particular, by Lemma (1.18), $P^{(M)}(T, x, \cdot)$ admits a density $p^{(M)}(T, x, \cdot) \in \hat{C}(R^M)$ for all $(T, x) \in (T_1, T_2) \times U$.

Given $\alpha \in (\mathscr{N})^N$, define $(T, x) \to \Lambda_\alpha(T, x) \in \mathscr{D}'(R^M)$ by

$$\Lambda_\alpha(T, x)(F) = D_x^\alpha \int_{R^M} F(y)\, P^{(M)}(T, x, dy).$$

Then

$$D_y^\beta(\Lambda_\alpha(T, x)) = \psi_{\alpha,\beta}(T, x, \cdot)\, p^{(M)}(T, x, \cdot) \in \bigcap_1^\infty L^q(R^M).$$

Hence, by the Sobolev embedding theory, $\Lambda_\alpha(T, x) = p_\alpha^{(M)}(T, x, \cdot) \in \hat{C}(R^M)$. Furthermore, one can easily check that for each $k \geqslant 0$ $\| p_\alpha^{(M)}(T, x, \cdot)\|_{C_b^k(R^M)}$ is uniformly bounded for (T, x) in compact subsets of $(T_1, T_2) \times U$.

Now choose $\rho \in C_0^\infty(R^M)$ so that $\int_{R^M} \rho(y)\, dy = 1$ and set $\rho_\varepsilon(y) = \varepsilon^{-M} \rho(y/\varepsilon)$, $\varepsilon > 0$. Define

$$u_\varepsilon(T, x, y) = \int_{R^M} \rho_\varepsilon(y - \xi)\, P^{(M)}(T, x, d\xi).$$

Then

$$\frac{\partial^n u_\varepsilon}{\partial T^n}(T, x, y) = L_x^n\, u_\varepsilon(T, x, y), \qquad n \geqslant 1$$

and for $(T, x) \in (T_1, T_2) \times U$:

$$D_x^\alpha D_y^\beta u_\varepsilon(T, x, y) = \int_{VR^M} \rho_\varepsilon(y - \xi)(D_\xi^\beta\, p_\alpha^{(M)})(T, x, \xi)\, d\xi$$

Combining these and letting $\varepsilon \searrow 0$, we conclude that $p^{(M)} \in C^\infty((T_1, T_2) \times U \times R^M)$. Q.E.D.

(7.8) *Remark.* There is no particular reason for our not letting σ and b depend on time as well as space. Indeed, with only minor modifications, the results of this section continue to hold for coefficients $\sigma(t, x)$ and $b(t, x)$ satisfying $\sup_{0 \leqslant t \leqslant T} \|D_x^\alpha \sigma(t, \cdot)\|_u \vee \|D_x^\alpha b(t, \cdot)\|_u < \infty$ for all $T > 0$ and $|\alpha| \geqslant 1$.

References

1. Bismut, J. M., Martingales, the Malliavin calculus and hypoellipticity under general Hörmander's conditions, to appear in the Proceedings of the L.M.S. Conference at Durham (July 1980).
2. Friedman, A. "Partial Differential Equations of the Parabolic Type," Prentice–Hall, Englewood Cliffs, N.J., 1964.
3. Gihman, I. I., and Skorohod, A. V., "Stochastic Differential Equations," Springer-Verlag, Berlin/Heidelberg/New York, 1972.
4. Haussmann, U., Functionals of Itô processes as stochastic integrals, *SIAM J. Control Optim.* **16** (1978), 252–269.
5. Holley, R., and Stroock, D., Diffusions on an infinite dimensional torus, *J. Funct. Anal.* **42** (1981), 29–63.
6. Ikeda, N. and Watenabe, S., "Stochastic Differential Equations and Diffusion Processes," in press.
7. Kunita, H., On the decomposition of solutions of stochastic differential equations, to appear in Proceedings of L.M.S. conference at Durham (July 1980).
8. Malliavin, P., Stochastic calculus of variations and hypoelliptic operators, *in* "Proceedings International Conf. on Stochastic Differential Equations at Kyoto (1976)," pp. 195–263, Kinokianiya, Tokyo and New York, Wiley, 1978.
9. Malliavin, P., C^k-hypoellipticity with degeneracy, "Stochastic Analysis" (A. Friedman and M. Pinsky,, Eds.), pp. 199–214, Wiley, New York/London, 1978.
10. Michel, D., Régularité des lois conditionnelles en théorie du filtrage non linéaire et calcul des variations stochastique *J. Funct. Anal.* **41**, (1981), 8–36.
11. Shigekawa, I., Derivatives of Wiener functionals and absolute continuity of induced measures, *J. Math. Kyoto Univ.* **20**, No. 2 (1980), 263–289.
12. Stroock, D., The Malliavin calculus and its application to second order parabolic differential equations, I, *Math. Systems Theory* **14** (1981), 25–65.
13. Stroock, D., and Varadhan, S. R. S., On degenerate elliptic–parabolic operator of second order and their associated diffusions, *Comm. Pure Appl. Math. XXV* **25** (1972), 651–713.
14. Stroock, D., "Topics in Stochastic Differential Equations," in press. Tata Inst. Lecture Notes series, Springer-Verlag, Berlin/New York.

Printed in Belgium